VOL. 9 No. 10, pp. 289-320 1947/1948

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS.

THE METAL-DIAZONIUM SYSTEM FOR PHOTOGRAPHIC REPRODUCTION

by R. J. H. ALINK, C. J. DIPPEL and K. J. KEUNING.

773.79

The metal-diazonium system for photographic reproduction, which has been developed in Philips' laboratory in Eindhoven in the course of the last few years, is based on the discovery that when a solution of a diazonium compound and a metal salt, say mercurous nitrate, is exposed to light, atomic metal — *in casu* mercury — is separated. The "latent" mercury image thus obtained can be transformed by physical development into a silver image and intensified. The light-sensitive system is obtained in the form of a film or sheet by impregnating a suitable carrier, say a strip of cellophane 40 μ thick, in a homogeneous solution of the said materials. The metal-diazonium system possesses an extremely high resolving power (> 1000 lines/mm) and allows of working with a very high gamma (6—8) whilst on the other hand also low gammas (1—2) can easily be obtained by varying external factors, *viz.* the moisture content or the intensity of exposure. The light-sensitivity of the system is in cellophane several times (in paper some tens of times) greater than that of the usual diazotype printing papers. This system, which was originally intended only for producing distortion-free copies of Philips-Miller sound film, lends itself excellently, *inter alia*, for the copying of picture-sound films, thanks to the external variability of the gamma. The impregnating of the cellophane base takes place on printing machines designed for the purpose, whilst at the same times these machines are fitted with a device for regulating the moisture content of the base. The good photographic properties of the system and the very low cost of materials open great prospects for its application on a large scale in all sorts of fields. In addition to the sound film and the picture-sound film also the field of micro and macro documentation is regarded as an important domain for the application of the system.

The metal-diazonium system is a new light-sensitive system that has been worked out by a group of scientists in the Philips Laboratory at Eindhoven. It possesses a number of remarkable properties making it exceptionally suitable for the photographic reproduction of pictures as well as of sound. Something has already been said about this system in a previous article ¹), where it was shown by a comparison with the usual methods of reproduction what place the new system could occupy.

Here we shall give a more detailed description of the fundamental principles of the metal-diazonium system and after going more deeply into its properties show how it is realized and practised, then concluding by dealing with a number of perspectives for the application of this new reproduction material.

Principle of the system

The light-sensitive material

Diazonium salts have the general chemical formula

$$[\mathrm{R{-}N{\equiv}N}]^+\mathrm{X}^-$$

in which R is an aromatic radical and X some anion. These compounds have been used for quite a time already for photographic reproduction methods. They are most familiar and most widely used in diazotype, a light-printing process for the multiplication of technical drawings and suchlike (tracings) made on transparent material.

Just as with all reproduction methods based on diazonium compounds, the fundamentals of this light-printing process lie in the following properties:

---

¹) C. J. Dippel and K. J. Keuning, Problems in Photographic Reproduction, in particular of Sound-films, Philips Techn. Rev. 9, 65-72, 1947 (No. 3).

a) Coupled with certain phenols or amines, the diazonium salts form what are known as azo-dyestuffs.

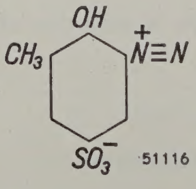
b) When a diazonium salt is exposed to light in the presence of water (water vapour) dissociation takes place and nitrogen is released. Schematically the reaction takes place as follows:

$$RN_2X + H_2O + h\nu \rightarrow \text{l.d.p.} + N_2 + HX \ . \quad (1)$$

The light-decomposition product (l.d.p.) thus formed is no longer capable of forming a dyestuff. If, therefore, one exposes a "diazotype paper" on which a tracing is laid and then causes the reaction sub a) to take place, a dyestuff is formed only on those parts of the paper that have not been exposed to the light. In this way a positive copy is obtained direct from the tracing.

The difficulties and limitations referred to in the previous article ¹) as being inherent in photographic reproduction with the usual silver-bromide and silver-chloride systems exist in a still higher degree with the diazonium processes hitherto known. They have not met with any appreciable success, for instance, in the reproduction of picture-sound films.

The new photo-chemical system described here is based on the discovery that the light-decomposition products obtained from certain diazonium compounds according to equation (1) are capable of releasing the metal from suitable metal salts, for instance mercury from mercurous nitrate, gold from aurochloride, etc. As diazonium compound one may use for instance o-cresol-diazonium-sulphonic acid:



When an aqueous solution of this compound and, for example, mercurous nitrate, $Hg_2(NO_3)_2$, applied in a thin flat layer is exposed with light of a short wavelength (*e.g.* 3650 Å) one may observe under the microscope small drops of metallic mercury at the places struck by the light. Thus a faint "mercury picture" is formed which might be compared to the "latent" picture in the usual method of photography with silver halogenides. And here, too, the "latent" picture can be intensified and made durable by developing. Contrary to the custom with silver halogenide photography, however, a so-called physical developing process is applied with our system. In view of the important part this developing method plays in the metal-diazonium system a separate chapter will be devoted to it below.

In order to show properly the difference from the old diazonium processes the point is stressed that only the second of the two characteristic properties of diazonium salts mentioned above, *viz.* the light decomposition, is utilized in the metal-diazonium system, the other property — formation of dyestuffs — playing no useful part in our system. The diazonium compound left on the unexposed parts is removed, while the exposed parts turn black (separation of metal). Thus we get instead of a positive copy a negative copy, as is the case with silver halogenides.

As to the practical realization of the system we shall revert to this later, but it is necessary to say something here about the manner in which the thin layer of the light-sensitive material is obtained. In this respect the new system differs fundamentally from the silver halogenide systems. With the latter systems an "emulsion" (more correctly: a suspension) of the light-sensitive substance, *e.g.* crystalline silver bromide, is made in gelatine and after a complicated and most precise ripening process this emulsion is cast on a celluloid film or a glass plate. In our case, on the other hand, a homogeneous solution is made of the diazonium compound and the metal salt and a suitable carrier is saturated with it. As carrier one may use for instance paper or, as we have done in the most important applications, a transparent base of regenerated cellulose. This latter material (which is more commonly known under the name of cellophane and is widely used as packing material for shop goods) is used by us in the form of a reel of film 0.04 mm (0.0016 inch) thick. But in principle one may also use with our system a gelatinous layer on celluloid or glass. One may also use the cellulose acetate film (so-called safety film) commonly applied for sub-standard films and by saponification make the surface suitable to absorb the aqueous solution of our light-sensitive system.

Physical developing process

What takes place in the so-called physical developing process in photography may be resolved into a phenomenon often met with in nature and in technology and which might be described as follows. When particles from an over-saturated solution or vapour begin to precipitate they show a preference for places where certain "nuclei" are present. If these nuclei are distributed locally in such a way as to form a "picture", maybe so faint that the eye cannot see it, then the picture is intensified by this selective precipitation of particles and may thereby be made visible.

A familiar phenomenon known to everyone is the picture formed by condensation on a smudged glass window when the window is breathed upon or cold air blows upon it from one side, water vapour condensing on the traces of dirt acting as nuclei and making visible figures [2]). Another example that may be given is that of the Wilson camera, where droplets of water from an oversaturated vapour condense on electrically charged particles and thus make visible the path followed by an ionizing particle.

Turning particularly to photography we see that both chemical and physical developing methods are in use, the latter more particularly in cases where negatives have to be intensified. Let us first consider what ordinary chemical developing comprises. A silver bromide film or plate is placed in a developing bath containing a reducing agent, for instance metol or hydrochinon, in a generally alkaline solution. The grains of silver bromide, in which silver nuclei have been formed by the exposure (groups of say 100 atoms of metallic silver), are reduced by the developer entirely to silver, whilst the unexposed or too weakly exposed grains remain untouched. After this developing, as is known, the remaining silver bromide is removed with the aid of sodium thiosulphate (fixing).

If the picture thus obtained is not dense enough it can be intensified by physical development. The film is placed in a weak acid solution containing silver nitrate in addition to a reducing agent, for which metol or hydrochinon can again be used. In the solution the silver nitrate is gradually reduced to silver, the solution becoming oversaturated, as it were, with atomic silver, which precipitates preferably on those places where "silver nuclei" are already present, *i.e.* on the exposed parts of the film [3]).

It is perhaps of interest to note that this physical developing process can be applied also to the "latent", not yet chemically developed image on the exposed silver bromide film. One can start by fixing the exposed film, thus removing all the silver bromide and leaving only the silver nuclei formed by the exposure, the latter then acting as nuclei for the subsequent physical development as described above.

The oldest photographic process, daguerreotype, was likewise based upon a physical method of developing; an image was formed on an iodized silver plate and the exposed plate was treated with oversaturated mercury vapour. Mercury was thereby condensed on the exposed parts of the plate and the picture became visible.

After these examples the method of physical developing applied with the metal-diazonium system does not need much more explanation. The carrier, containing for instance o-cresol-diazonium-sulphonic acid and mercurous nitrate, is placed after exposure in a suitable solution of silver nitrate and a reducing agent. Reaction then takes place between the locally formed metallic mercury and the silver nitrate, the mercury dissolving and silver precipitating at those places, the latent mercury picture thus being transformed into a latent silver picture. Moreover, by reduction metallic silver is gradually formed in the solution. The local metal deposits already present act as nuclei upon which more and more silver is deposited as new silver is formed in the solution, thereby developing the picture.

One of the essential factors in this process is the rapidity with which the silver is formed in the solution by reduction. If this takes place very rapidly one gets instead of a selective deposit a more or less evenly spread deposit of silver, a phenomenon that can even be turned to advantage for making homogeneous silver mirrors. By giving the solution, for instance, a suitable degree of acidity one can regulate the speed of the spontaneous reduction of the silver nitrate and cause practically all the silver formed during the developing process to precipitate on the nuclei of the picture, to the exclusion of almost all undesired precipitation of silver on the unexposed parts. The negative may not be kept in the developing bath longer than the time taken for developing, because the developer is an unstable system and liable in time to cause a spontaneous flocculation of all the silver, which then precipitates anywhere.

We will not conclude this explanation of the principles of the metal-diazonium system without remarking that here only a very rough and greatly simplified representation of the processes has been given. To understand the connection between the photographic properties obtained and the numerous variable factors of the system it has been necessary to study deeply the mechanism of the light decomposing reaction [equation (1)] and of the developing process. It may be possible at a later date to go more deeply into the problems arising, some of which are still unsolved.

---

[2]) In the familiar game of writing or drawing on a glass window with a wet finger, when some days later the writing is made visible by breathing upon it, we have examples of negative pictures; the window is more or less uniformly covered with "dirt", which is removed by the finger (or at least partly so) or else a thin film of grease is left behind, so that less vapour is condensed on the writing than on the rest of the window.

[3]) In this case the physical developing process comprises also a chemical process. From this it appears that the name "physical development", which we use because it has already been introduced, does not express the essential difference from the actual chemical process of developing. The difference as we see it lies in the fact that in chemical development the metal from which the picture is built up is already present in the appointed place prior to developing in any form whatever, whereas in the physical development it is only brought into its appointed place by the developing process.

Properties of the metal-diazonium system

Resolving power

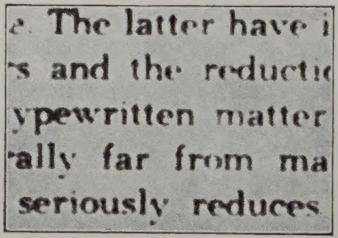
In the article previously quoted [1]) it was explained that for photographic reproduction, especially of picture-sound films, a high resolving power is favourable. However, there is a limitation in this respect, in that it tends to spoil the quality of reproduction (unsharp pictures, distortion of high frequencies), or, when trying to avoid this drawback, it necessitates a greater length of film (the picture may not be too small), the same applying for the speed of the film on account of the sound.

The positive films commonly used for copying picture-sound films have a resolving power of between 50 and 75 lines/mm, so that when projecting gratings finer than what corresponds to this number of lines per mm the lines run into each other and therefore can no longer be made visible separately. Some new Kodak films for special purposes have a resolving power as high as 160 lines/mm. *The metal-diazonium system is quite capable of resolving 1000 lines/mm.* It is even possible that its resolving power is still greater, for the optical means used in determining this power have themselves a limited resolving power, which in our case did not reach further than the said limit of 1000 lines/mm.

In other laboratories too, *i.a.* Kodak, it has for some time been possible to make reproductions with an extremely high resolving power, of the order of 1000 lines/mm [4]). Collodion plates or so-called Lippmann emulsions are used. Except for the already known application of wet collodion plates for making autotypes, as far as we know none of these processes has been suitably developed for general use.

The high resolving power of the metal-diazonium system is due for a large part to the fact already mentioned that the light-sensitive material is not used in the form of an emulsion but in that of a homogeneous solution; the unexposed system is quite free of grains, so that there is very little diffusion of light in the sensitized material. If a large quantity of light reaches one point then in the circle round about that point where undesired light reaches, owing to diffusion, the quantity of that light is extremely small. Moreover — and this is the second cause of the high resolving power — that circle (the diffusion halo) remains extremely small owing to the fact that with the chosen concentration of the chemical components in the light-sensitive solution the active light is very strongly absorbed. Consequently the diffused light does not reach far.

Owing to the strong absorption the direct light thrown upon the carrier upon exposure stays in the top layer. Consequently the metal forming the picture is limited to a thin layer. This is likewise of importance for the high resolving power, for it is not sufficient that details are recorded well-separated in the carrier — they must also be reproduced separately either for copying or for projection. Owing to the fact that one never uses perfectly parallel light beams, the thicker the picture layer the more details are lost [5]).



e The latter have i
s and the reductic
ypewritten matter
ally far from ma
seriously reduces

50928

Fig. 1. A reproduction, linearly enlarged about 250 ×, of a piece of a micro-document recorded with the metal-diazonium system in cellophane. The size of this piece in the micro-document was 0.012×0.017 cm and the height of the letters 12 microns. The good definition of the reproduced letters gives an idea of the exceptionally high resolving power of the system (> 1000 lines/mm).

Fig. 1, the magnified reproduction of a micro-document made with the new system, demonstrates the high resolving power.

---

[4]) See for instance J. Sci. Instr. **18**, 66-67, 1941, where an account is given of the "Kodak maximum resolution plate", which can resolve 600 or even 1200 lines/mm. Similar results have been reported by H. Frieser, Z. wiss. Phot. **40**, 132, 1941. For older methods see E. v. Angerer, Wissenschaftliche Photographie, Akad. Verl. Leipzig, 1931, p. 136 *et seq.*

[5]) Also in the old diazotype processes based on diazonium compounds the material is free of grains, but the resolving power is not very high, because, *i.a.*, the picture is rather thick: owing to the relatively small absorption a thick layer of dyestuff is necessary for adequate "density". Furthermore the formation of the dyestuff is a relatively slow reaction, so that after the exposure a noticeable diffusion takes place before the (fixed) molecules of the dyestuff are formed.

The gamma value

Every photographic picture, be it positive or negative, has a certain characteristic density curve which indicates the density obtained for any exposure $E = I \cdot t$ (I luminous intensity, t exposure time). The density D 6) is usually plotted as a function of log E; see for instance *fig. 2*. For many

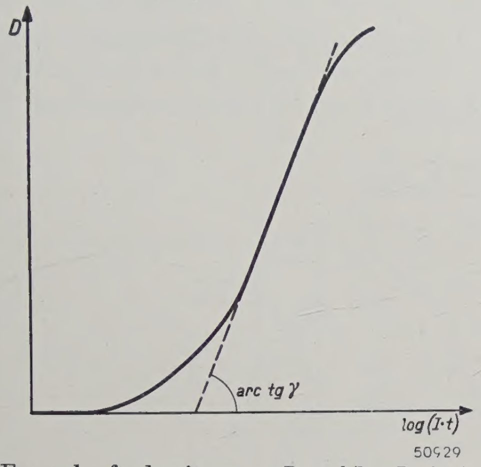


Fig. 2. Example of a density curve $D = f$ (log $I \cdot t$) of a photographic picture. The maximum slope of the curve lying in the practically rectilinear part in the middle is the gamma.

light-sensitive materials the density curve shows a more or less extended, practically rectilinear part, where at the same time the slope of the curve is greatest. The whole curve is in fact characterised by this slope, the gamma, because this is decisive for the gradation in the reproduced picture (reproduction of the shades of brightness). It is well-known that with the usual silver halogenide systems the value of the gamma depends largely upon the emulsion and is further particularly determined by the conditions of developing (temperature and composition of the developing bath, time taken in developing). Common values in practice vary from 0.5—2.5.

When applied in a suitable manner, as will be defined below, the metal-diazonium system has very *much higher gammas, e.g. 6—8.* In the previous article 1) it has been explained what important advantages this offers for instance for sound reproduction. Since with a high gamma a relatively small reduction in the exposure intensity is sufficient to bring about a transition from the "greatest density" to the "smallest density", only the innermost part of the circle diffusion halo already referred to is noticeably blackened. Thus the high gamma promotes a high resolving power and therefore in the case of sound reproduction promotes good reproduction of the high frequencies. The high gamma yields particular advantages in the copying of Philips-Miller film, as we have seen in the previous article; the "lens effect" arising in this process is rendered harmless without any other measures being necessary 7).

External variability of the gamma

In sound reproduction by the amplitude system advantage can safely be taken of a high gamma, because in principle we have only to do with two densities, a veıy low one in the transparent sound track and a very high one for the rest of the film. There are all sorts of other applications where the same holds, *e.g.* in micro and macro documentation, to be discussed below. In picture reproduction, on the other hand, a whole series of shades of brightness (half tones) has to be reproduced by corresponding, continuously varying densities. To attain this it is necessary to satisfy the Goldberg condition 8), as a result of which in picture reproduction one has to work with relatively low gammas, *e.g.* 1.5—2.5.

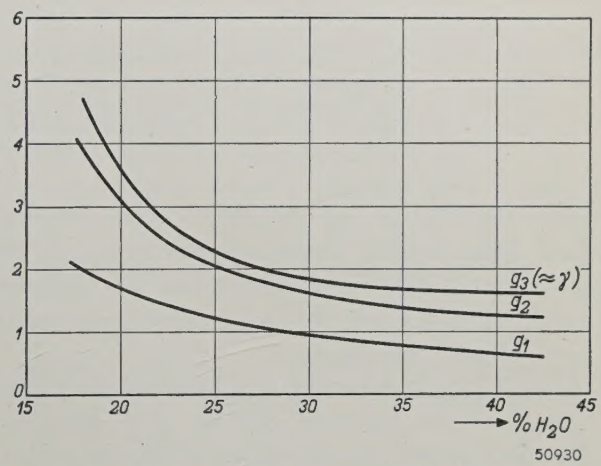


Fig. 3. Effect of the moisture content of the cellophane film upon the gamma of the mercury-diazonium system in the film for a certain concentration of the ingredients and a certain manner of exposure and developing. The lines plotted represent the average slopes g_1, g_2, g_3 of the density curve in three different density areas; g_3 is practically equal to the maximum slope γ.

---

6) Defined as $D = \log i_0/i$, where i is the portion of an incident quantity of light i_0 that the blackened plate or film allows to pass through.

7) Also the so-called wedge effect with thePhilips-Miller film is rendered harmless by the high gamma. When the transparent sound track in a "Philimil" film is cut with the wedge-shaped chisel, which is a characteristic feature of the Philips-Miller system, an oblique edge is left on the top layer, and when copying this results in a gradual transition from the maximum to the minimum density. In itself this does no harm, but aberrations arise if there is any variation in the thickness or the density of the covering layer. The higher the gamma, the narrower this edge transition shows on the copy and the smaller the aberration.

8) See the explanation in the article quoted in footnote 1.

Now such gammas and still lower values can be obtained quite easily with the metal-diazonium system. This is possible not only by a suitable choice of the composition of the system but also by making use of the following important feature: *the gamma of a metal-diazonium-cellophane system can be greatly influenced by the moisture content of the film during exposure and likewise by the duration of the exposure.* The very high gammas referred to under the previous heading occur when the cellophane film is dry, that is to say when its moisture content is not more than say 15% by weight, and when the film is exposed with a great intensity (and corresponding short exposure time). If the moisture content is increased to 25—30% by weight then one gets the low gammas required for good picture reproduction.

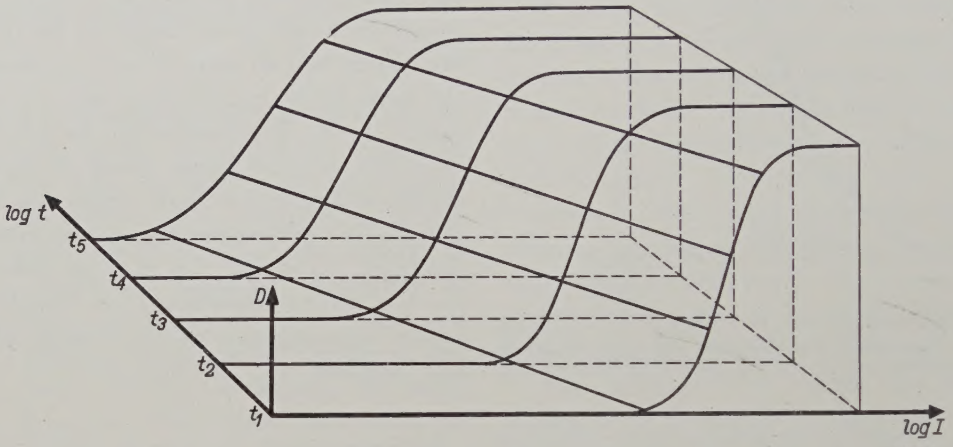


Fig. 4. The gamma of the metal-diazonium system also depends, *i.a.*, upon the intensity of exposure I. Therefore the density D is not fully determined by the product $I \cdot t$: the photographic behaviour of the system has to be described by a solid density plane $D = f\ (I, t)$. This plane is drawn here for given concentrations, moisture content, developing, etc. in perspective. Log I and log t are plotted as independent variables.

Fig. 3 indicates the relation between the gamma value and the moisture content of a cellophane film for certain concentrations of the ingredients and under certain conditions of exposure, developing, etc. There the following refinement has been applied. Instead of showing the gamma indicating the maximum slope of the density curve, the values g_1, g_2, g_3 have been plotted of the average slope that the density curve assumes in three consecutive density intervals of the most importance in practice, *viz.* between $D = 0.05$ and 0.5 (high lights); $D = 0.5$ and 1.0 (intermediate tones); $D = 1.0$ and 1.5 (shadows). The value g_3 is generally practically equal to the maximum slope defined as γ.

This defined description of the gradation, often applied with silver halogenide systems and sometimes even extended to five intervals, is desired when one has a group of density curves which do not all have a rectilinear part with maximum slope in the density area required. This is also frequently the case with the metal-diazonium-cellophane system, with which it is possible to get a large variety of density curves under the influence of the numerous variables.

The fact that the gamma, or in other words the density curve, depends upon the duration of exposure t means that the density D is no longer a function of the product $I \cdot t$ but of I and t separately. Therefore instead of a two-dimensional density curve, to describe fully the photographic behaviour of the system we need a solid figure in which the curved plane $D = f\ (I, t)$ is represented. Such a plane is drawn in perspective in *fig. 4.* It must be borne in mind that by varying the moisture content, the conditions of developing, etc. a different plane is obtained every time. The "density curve" of a picture taken with an exposure time t_1 is the cross section of the density plane parallel to the D and I axes which is intersected by the t axis at t_1. From fig. 4 it is clearly seen that for different values of t one obtains density curves with a different slope (gamma). A number of these curves are drawn in *fig. 5*, whilst in *fig. 6* the slopes g_1, g_2, g_3 ($\approx\gamma$) of these curves are plotted as a function of t.

From this it also follows that the relation between log I and log t for a constant density D (horizontal cross sections of the curved plane at different heights; see fig. 4) is given by lines the slope of which must gradually change with varying D. Whereas in the simplest case, which was assumed in fig. 2, D was only a function of $I \cdot t$, so that for the given D we had $\log I + p \log t =$ constant with $p = 1$, in our case as a rule $p \neq 1$. This is the familiar Schwarzschild effect. Moreover, according to the foregoing, in our case the Schwarzschild exponent p (slope of the log I-log t lines in the area where these may be regarded as straight) depends also upon the density.

Figs. 3 and 6 give a clear picture of the great variability of the gamma in the metal-diazonium system. Now it is important to note that this is an *external variability*: on one and the same material and with one developing process we can make copies with a high gamma and with a low gamma. This creates not only the possibility already mentioned of copying both sound and pictures each with the most favourable gamma, but also an entirely new possibility of copying sound and picture side by side on one film and developing them together without necessitating a compromise in the gamma such as is characteristic for the present picture-sound film technique (see the previous article [1])). All that is necessary is either to select the exposure intensities separately for the copying of the pictures and for the copying of the sound track, or else to vary the moisture content of the cellophane band between the two places where in the printing machine first the picture is copied and farther on the sound track is copied (it may be that both measures have to be applied together).

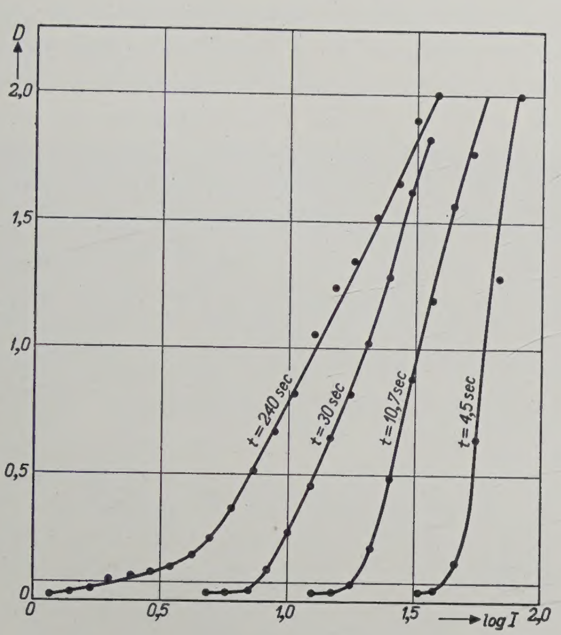


Fig. 5. From the density plane drawn in fig. 4 it is possible to find for any exposure time t the density curve $D = f(I)$ of the picture obtained with that exposure time, by taking a cross section of the plane perpendicular to the t-axis at the level of that time. Such cross sections are represented here for various exposure times. It is seen that curves are obtained with different slopes (different gammas).

In a similar manner we can also print on one film the pictures of different scenes with a different gamma while requiring only one developing process! This offers the possibility of correcting any exposure or developing variations in the negatives.

Light sensitivity

The light sensitivity of the metal-diazonium system depends not only upon the choice of the ingredients, etc. but also varies with different carriers. When cellophane is used the sensitivity may be several times greater than that of the known diazotype papers. Yet with paper as a carrier the sensitivity of the system is greater by a factor of 10. For instance with mercury as metal the metal-diazonium paper can easily be made 20—25 times as sensitive as the positive diazotype papers, a gain which just makes it possible to produce enlargements by the light-printing process (see the last chapter). Still this is a factor 10^4 below the sensitivity of silver bromide enlargement papers.

There is, therefore, no question that the metal-diazonium system as at present developed could compete with the materials commonly used for photographic recording. As a matter of fact the spectral zone in which the metal-diazonium system is sensitive is too limited for this purpose.. The sensitivity of the mercury-diazonium system lies mainly in the near ultra-violet with a maximum in the vicinity of 3900 Å and not extending beyond the bluish green (about 5000 Å).

For application as reproduction material, however, in most cases neither the low sensitivity nor the limitation of the spectral area constitutes any objection. One only needs to use for the copying process a light source possessing great luminous intensity in the said range of the ultra-violet. Super-high-pressure mercury lamps with water-cooling are excellently suited for this purpose. With these lamps the intensity of light that can be reached when exposing the film is so great that a film can be copied at fairly great speeds, for instance 20 meters per minute.

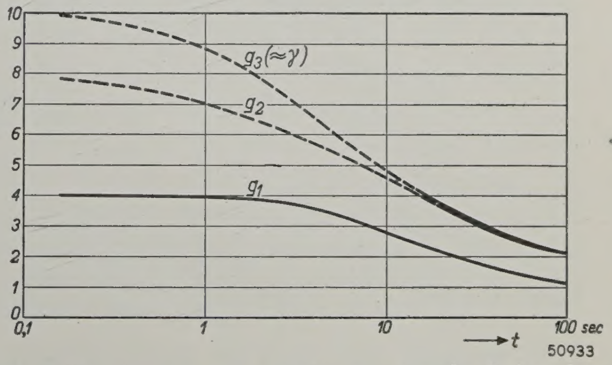


Fig. 6. The slopes g_1, g_2, g_3 of the curves of fig. 5 plotted as functions of the exposure time t.

An important advantage of the limited spectral sensitivity lies in the fact that the whole working of the system can take place under a bright sodium light, since in the wavelength range of the sodium lines (5890 Å) the sensitivity is practically nil.

A great deal of work has been done in this laboratory in respect to the question as to what determines the sensitivity of the metal-diazonium system. In the particular case where mercury is employed as a metal it was demonstrated that the process of the formation of the latent mercury picture has a quanta yield of about 50%; for an average of two exposure quanta one atom of metallic mercury is formed. This means that the "primary" sensitivity is of the same order as that of the silver halogenide systems. According to the conclusions provisionally reached from an investigation into the highly complicated mechanism, the cause of the so much smaller resulting sensitiveness of our system is to be sought rather in the further history of the nuclei to be developed; the metallic atoms combine to form larger particles. As already stated, under a very strong magnification (about 800 ×) these particles may be seen in the latent picture. Thus in this stage the metal is rather coarsely dispersed and in the physical developing process the resultant density is all the less according as the (given) quantity of metal of the latent picture is more coarsely distributed.

The influence of the moisture content and the intensity of the exposure upon the gamma is also closely related to the history of the primary metallic atoms and of the metallic nuclei prior to developing.

Durability

When talking of durability in the case of the metal-diazonium system we have to differentiate between the exposed and the unexposed state. The unexposed system appears to have as yet too little durability for a sensitized cellophane film, for instance, to be kept in stock until one has need of it. For the most important applications that we have in mind, however, this forms no serious objection, as will be made clear below.

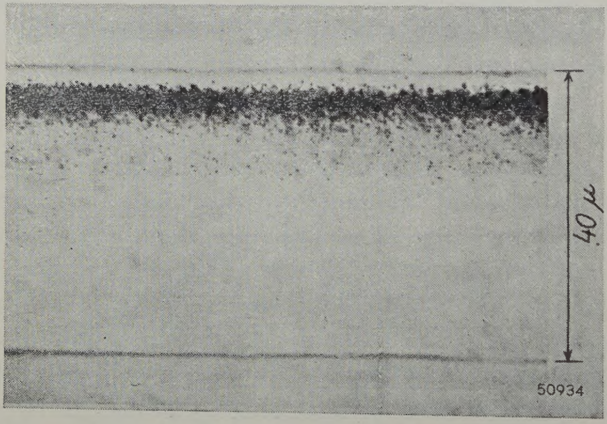


Fig. 7. Microtome cross section of a sensitized cellophane film 40 microns thick after exposure and developing. The particles of silver from which the picture is built up lie in a thin layer a few microns below the surface.

After exposure and developing one has a picture that will keep practically indefinitely: as already mentioned above, the definitive picture consists of metallic silver, just as is the case with silver halogenide systems, and thus is perfectly proof against atmospheric influences and light (such contrary to the dyestuff pictures of the old diazotype papers). Furthermore, the picture on a cellophane film is protected in a peculiar manner against mechanical damage: when a microtome section of such a film is cut and examined under the microscope it will be seen that the extremely thin layer containing the metallic silver of the image (see above) does not lie on the surface of the film but a few microns below it; see *fig.* 7. We cannot go into the explanation of this here, but it has the welcome practical advantage that the actual picture is protected against scratching etc. by the thin layer of clear cellophane covering it.

Economy

It is to be expected that the cost of the metal-diazonium system will be relatively low, a factor that will prove to be of great weight for various applications. How is it that the system turns out to be so economical? In the first place there is the choice of the carrier, cellophane being the cheapest material imaginable for this purpose. Another important factor is the limited consumption of silver, it being a characteristic of the physical developing method that the sensitive material itself need not contain any silver; the silver comes from the developer and is added to the negative, and the developing can be so controlled that not much more silver need be used than is necessary for building up the ultimate silver picture. With the usual silver halogenide systems, on the other hand, a very high percentage of silver is wasted; it comes out of the exposed film in the fixing bath and cannot be recovered except at considerable expense.

Of further importance is the very simple method of manufacture, as will be explained under the next heading.

Yet another advantage to be mentioned in connection with the economy of the metal-diazonium cellophane system is the saving in volume and the attendant ease of storage and transport. A cellophane film 40 microns thick and say 300 meters long winds up into a reel about 13 cm in diameter, whereas a normal celluloid film of 300 meters length forms a reel 26 cm in diameter.

Realization of the metal-diazonium system

The metal-diazonium system can be realized in quite different ways according to the use intended and the carrier employed. As a typical example we will consider here the application of the system for the copying of picture-sound films with cellophane as the carrier.

Just as is the case with the common silver bromide films, so with the cellophane film the actual copying process takes place on a machine where both the negative and the positive films are caused to pass along under a lamp simultaneously. Now we have already said that the cellophane film is sensitized by impregnating it in its entirety with the solution containing the light-sensitive system. Further it has been stated that the cellophane film sensitized in this manner has only a limited durability. The difficulties that this might involve have now been overcome in a very simple way by *combining the sensitizing process with the printing process*, the impregnating of the cellophane film and the exposure taking place on the same machine in succession.

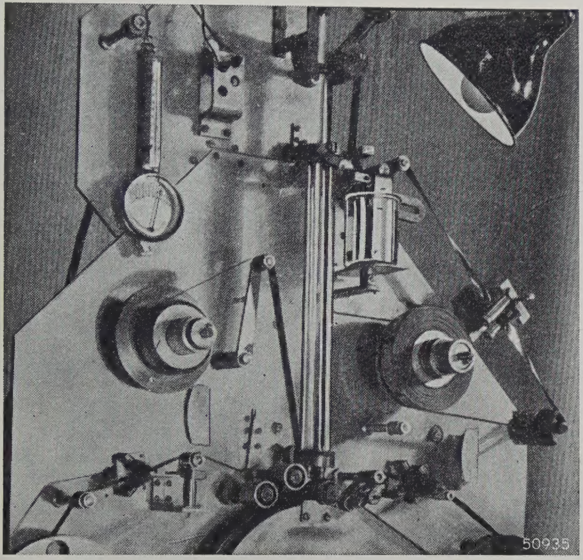


Fig. 8. Printing machine on which a cellophane film is first impregnated in the light-sensitive solution, then dried to the desired moisture content and after that exposed. The film is fed in from the right, impregnated in the bath at the top on the right, dried in the vertical tube and printed on the drum in the middle at the bottom of the photo, where it is brought into contact with the original film to be copied, the two films passing underneath the lamp simultaneously. (For drying the film when running at a high speed several tubes were used; with the latest machine drying is done by high-frequency heating.)

The fact that we have here an extremely simple and economical *modus operandi* is quite evident when comparing it with the manufacture of silver bromide film, where the preparation of the carrier, the preparation and ripening of the emulsion, the casting of the emulsion on the celluloid film and later the copying are all done in separate departments.

The first mentioned method has also been found to improve considerably the reproducibility of the properties of the film.

Between the impregnating and the exposing of the cellophane film this has to be dried to a moisture content corresponding to the desired gamma. Further, it must be possible to reduce still further the moisture content between the exposure of the picture and that of the sound track on the film if such should be desired. This can be done by passing the film through a tube with conditioned air, as seen in *fig. 8*, or by means of high-frequency heating.

Possibilities of application

The system is still too young to allow of any data being given as to its applications, but the experience so far gained with it opens such interesting perspectives that a brief outline of some of its possible uses may well be given here.

Sound film

A stereophonic sound track has been made on 7 mm cellophane films by copying a stereophonic Philips-Miller film [9]). Thanks to the sharp definition of the mechanically recorded original and the high resolving power of the copying material an exceptionally good quality of sound is obtained, in respect to the reproduction of the high frequencies and the absence of non-linear distortion (see the article quoted in footnote [1])). This good quality of reproduction together with the enhanced "naturalness" obtained by stereophony seem to us to constitute the requisites for imparting to "mechanical" music the original musical character.

Partly by reason of the low cost of the reproduction method, it is in principle possible that this ideal method of reproducing music will come within the reach of everyone for use in the home. An attraction of the cellophane films used for this method is that a music film with a playing time of one hour forms a reel no more than 18 cm in diameter (playing speed 32 cm per sec); see *fig. 9*. This is due on the one hand to the extreme thinness of the cellophane film (40 microns) and on the other hand to the fact that it is possible to print on a 7 mm film two stereophonic sound tracks (thus in all 4 tracks).

---

[9]) K. de Boer, Stereophonic Recording on Philips-Miller Film, Philips Techn. Rev. **6**, 80-84, 1941.

A piece of music of one hour can be reproduced without any of the interruptions that are unavoidable with the gramophone even when using an automatic record-changer.

Picture-sound film

Apart from the cinema there is a wide field of possibilities awaiting the "talkies". It could be used on a large scale for entertainment in the home, for educational purposes in schools, for advertising, etc. provided it is cheap and of good quality. The 8 mm film, which would seem to lend itself best to this purpose, has not yet been widely used because it is still too expensive and owing to the limited resolving power of the common emulsions the pictures are not sharp enough; furthermore it does not leave any room for the sound track. Though the sond track is applied on a 16 mm film the quality of the sound reproduction is not all one would desire.

Now the metal-diazonium system as a copying material presents a situation that is more favourable in many respects: apart from the possibility of making a sound track even on an 8 mm film, on a 16 mm film a better sound quality can be obtained than has hitherto been possible, partly due to the high resolving power and partly by reason of the variability of the gamma. Both of these factors also help in improving the picture quality — as is clearly noticeable on a contact print of a very fine-grained film, *e.g.* Isopan FF — although a limit is set to the sharpness of the picture by the limited resolving power of the original film (about 55—75 lines/mm) and possibly of the optical system of the camera. Finally the film could be cheaper.

The same considerations apply for the copying of standard 35 mm films on the new system.

For playing the very thin cellophane films special projectors are required. *Fig. 10* shows a model of a home cinema equipped with such a projector. If normal projectors are still to be used then the metal-diazonium system will have to be applied in or on a thicker carrier (say 130 microns thick), but then of course one loses the advantage of the great compactness of film reels with a long playing time.

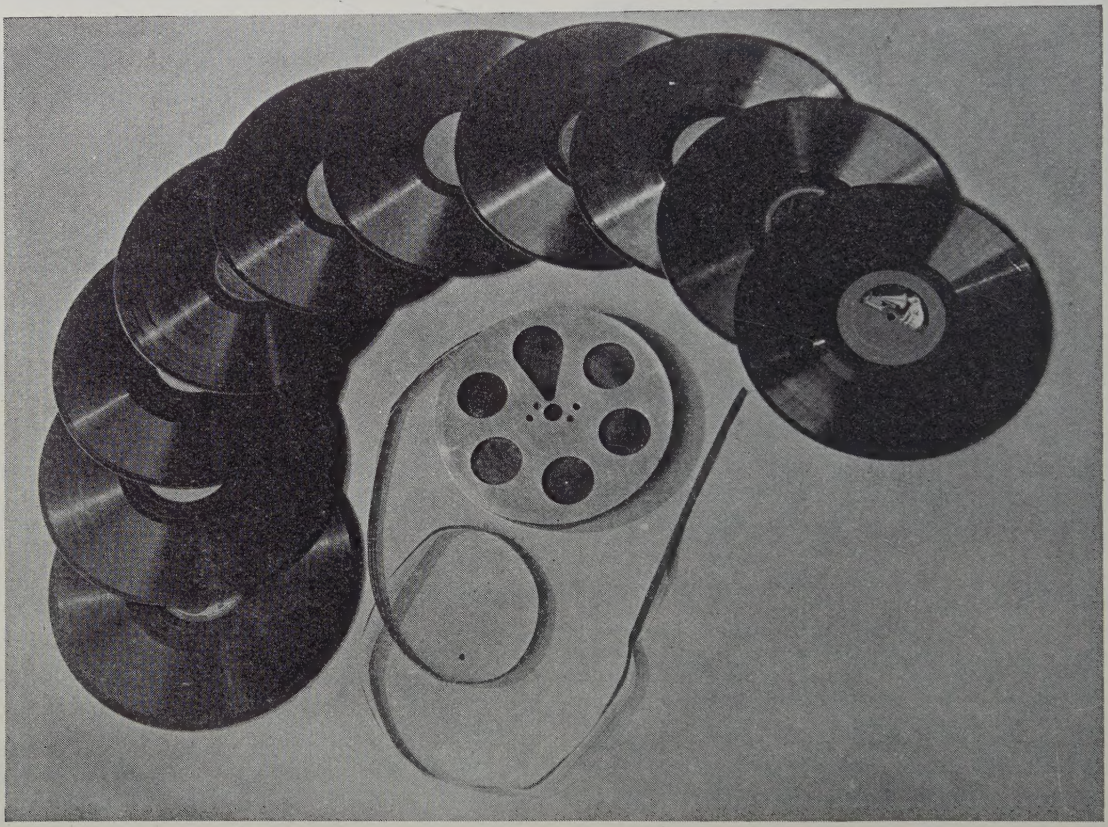


Fig. 9. An illustration of the compactness of a recording of music on cellophane film. Music that takes one hour to play can be recorded stereophonically on a film reel of the size of that shown in the illustration. For the same playing time (without stereophony!) 10 gramophone records of 25 cm diameter are required.

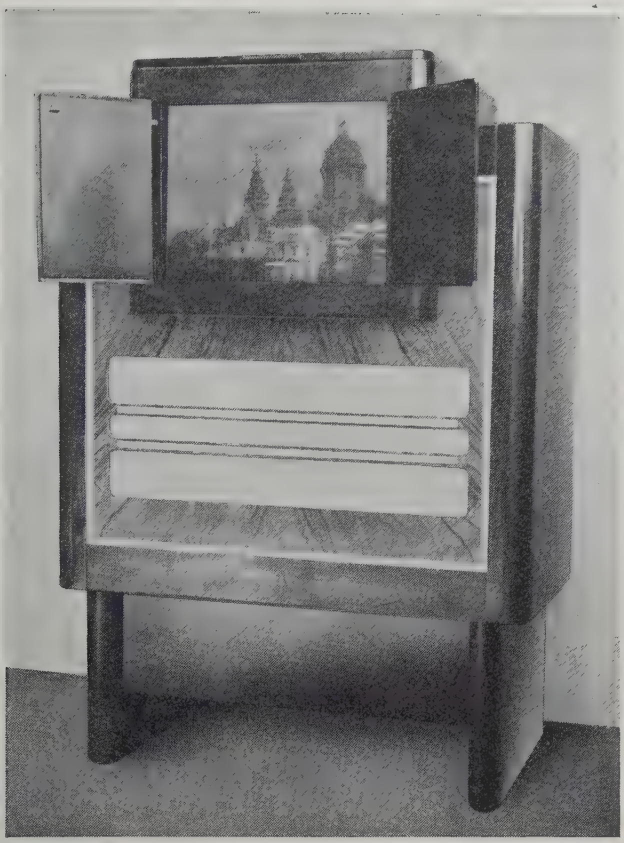


Fig. 10. Model of a home cinema fitted with a special projector (mounted in the cabinet) for cellophane film. The projected picture satisfies high requirements regarding sharpness as well as gradation.

Micro-documentation

Micro-documentation is a comparatively young branch of the technique of reproduction, but it seems that a surprising development may be expected in this very direction.

The name itself already expresses its meaning: the recording of documents on a very small scale. The need for this may be due to various reasons. In some cases it is resorted to because the documents are so bulky or would become so bulky as to constitute a problem for their filing, storage, handling or transportation; examples are the catalogues of large libraries, card index systems of large telephone exchanges or of the registers of births, deaths and marriages of large cities, etc. There are other cases where micro-documentation is applied because reproduction on the normal scale is too expensive, as for instance the copying of publications by libraries, or for the air-mailing of documents where weight is a big consideration, etc.

Obviously the higher the resolving power of the material used for the photographic reproduction, the smaller the size of the reduced document. With the extremely high resolving power of our metal-diazonium system it is possible to make a perfect record of a whole page of printing of the size of this journal on an area of 0.6 by 0.9 mm, the height of the letters being about 12 microns.

It cannot be predicted whether one will ever go as far as such an extremely small size with the present stage of development of micro-documentation. For the present the sizes of for instance 5 × 7 or 2 × 3 mm seem desirable.

If one keeps to the larger dimensions of say 5 × 7 mm per page then the high resolving power of the metal-diazonium system is not utilized to its fullest extent, but even so this material will prove to be of great advantage owing to its low cost. Large tabular works, encyclopedia, etc., which can now practically only be consulted in libraries, could be reproduced on such a small scale on this inexpensive material as to be brought within the reach of anyone having occasion to read such works from time to time. Micro-reproductions can be read with a simple reading apparatus. The saving in volume is astonishing, even with the relatively

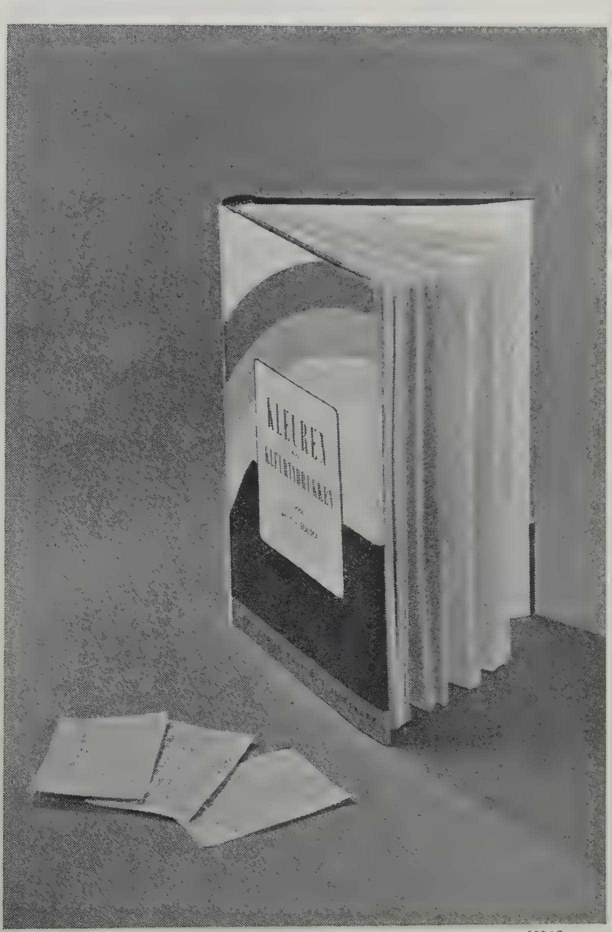


Fig. 11. Three small sheets of metal-diazonium paper on which the complete contents of a book of 330 pages have been reduced. Each page of the book is reduced on the paper to a size of 5 × 7 mm. The reproduced book is quite legible with a simple reading apparatus.

large size of 5 × 7 mm, for a series of large books totalling 10 000 pages can be reduced to a pocket-size booklet of 100 pages. Three of such pages, compared with the original normal book, are shown in *fig. 11.*

Macro-documentation

As the last field of application for the metal-diazonium system we would mention that of macro-documentation, which has already become of common usage in the form of diazotype and blue-printing and photocopying directly on a legible scale. Owing to the nature of the documents to which this process is applied (drawings, specifications, etc.), for this purpose the metal-diazonium system would be employed in paper. The paper is sensitized on both sides, so that it can be used on both sides for different copies; it does not curl up, and on account of its two-sided use it sometimes means a considerable saving in volume. Furthermore the image (a silver picture) gives a very rich contrast, with a pleasant tone (neutral grey), and is quite stable, properties which are often so lacking in diazotype and blue-printing processes. Of particular importance, however, is the possibility of making enlargements with metal-diazonium paper with exposure times that are practicable (for instance 1-5 seconds, depending upon the size). Thus we get a very efficient working method: wherever such is desired on account of frequent use, documents recorded on a micro-film can be enlarged again on metal-diazonium paper.

OPTICAL ABERRATIONS IN LENS AND MIRROR SYSTEMS

by W. de GROOT.

535.317.6

When dealing with technical-optical problems, such as X-ray screen photography and television projection, the need is felt of an optical system with large aperture. Apart from the special lens systems developed for this purpose, Schmidt's mirror system, consisting of an aspherical correction plate, demands attention.

In this article the principal aberrations (so-called third order aberrations) of optical systems in general and of the spherical mirror in particular are discussed, and it is shown how with one exception (the curvature of field) all the third order aberrations of a spherical mirror can be eliminated with the aid of a correction plate.

Introduction

In the field of illumination, photography, picture projection, etc. innumerable technical problems arise where the need of an optical system with large aperture is felt. Certain special applications, as for instance X-ray screen photography and television projection, have emphasized the need of such a system in the Philips laboratory.

Complicated lens systems with large relative aperture have already been developed for this purpose, but now a second solution, the application of mirror systems, is demanding more and more attention.

Following the invention by Schmidt, who for astronomical purposes fitted a spherical mirror with an aspherical correction plate eliminating the most important aberrations, Philips have developed a mirror system with a correction plate which can be applied for the purpose mentioned above and which is simple to construct.

The object of this article is to examine more closely the aberrations of optical systems in general and of spherical mirrors in particular. The properties and construction of the correction plate will be discussed in a subsequent article.

Causes of the aberrations

It is already known that, with the aid of a lens or system of lenses whose limiting surfaces are centred spherical surfaces, and also with aspherical concave mirrors, waves of light emitted from a point source can be approximately concentrated at another point (the image point). The fact that this imaging is imperfect is due to two causes.

In the first place after refraction or reflection from a convex surface the spherical light waves emitted from the point source will generally be no longer truly spherical and consequently will not converge into one point.

The wave surfaces are always limited by the edges of the lenses or mirrors and often also by a diaphragm expressly fitted for the purpose. As a rule the aberrations due to the aspherical shape of the wave surfaces become greater the larger the diameter of the limiting apertures.

On the other hand, as a result of this limitation so-called diffraction phenomena arise which also cause lack of definition in the image. The effect of these phenomena is the greater, the smaller the diameter of the limiting aperture relative to the wavelength of the light used. It therefore depends upon the circumstances whether the first or the second cause predominates in the resulting lack of definition. With photographic lenses and with the mirror systems that will be discussed in this article one may, as a rule, ignore the diffraction effects. It is then permissible to substitute for the conception of light waves that of light rays, that is to say very narrow pencils of light which pass through the system independently of each other and are everywhere at right-angles to the wave surfaces. In other words one is satisfied with the approximation offered by geometrical optics.

First-order aberrations

Fig. 1 is a representation of the well-known elementary construction of the passage of rays through an optical system with rotational symmetry. The rays emitted from point P at a distance x_1 in front of the first focus F_1 and a distance y_1 below the axis converge at a point Q at a distance x_2 behind the second focus F_2 and a distance y_2 above the axis, where the equations

$$y_2 : y_1 = f : x_1 = x_2 : f = v.$$

apply. Thus the plane through P at right-angles to the axis is uniformly projected with a constant linear magnification v upon a plane through Q at right-angles to the axis. This is only correct to a sufficient approximation when all the rays pass so close to the axis ("paraxial") that the angles formed by all parts of the rays with the axis are so small that their sines and tangents may be interchanged. The quantity v is called the paraxial enlargement.

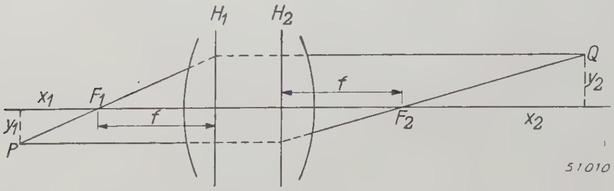


Fig. 1. Elementary construction of the image Q of a point source P in the case of paraxial passage of the rays. H_1 and H_2 represent the principal planes of the system; the focal distance is f.

Even in the simple case of the paraxial passage of rays represented above one may speak of aberrations. Let us suppose (*fig. 2*) that a narrow pencil of rays converges at a point Q in the plane V and a distance y_0 from the axis. If the rays are collected in another plane W not coinciding with V then instead of a point of light a small patch is observed in this plane, displaced with respect to the point Q. In particular, a ray intersecting a plane U at a point D having as coordinates $y = H \cos \varphi$, $z = H \sin \varphi$, will strike the plane W at a point having the coordinates

$$y = y_0 + a_1 H \cos \varphi + a_2 y_0, \quad . \quad . \quad \text{(1a)}$$

$$z = a_1 H \sin \varphi, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(1b)}$$

in which a_1 and a_2 are proportional to the distance d between V and W.

Let φ assume all values between 0 and $2\,\pi$, then the point y_u, z_u describes a circle with radius H in the plane U. The ray DQ thus describes a conic surface. Next let H assume all the values between zero and a maximum value H_1. Then we get a cone entirely filled with rays. One may also imagine this having been brought about by the introduction in the plane U of a circular diaphragm with radius H_1. We shall therefore call the plane U the diaphragm plane.

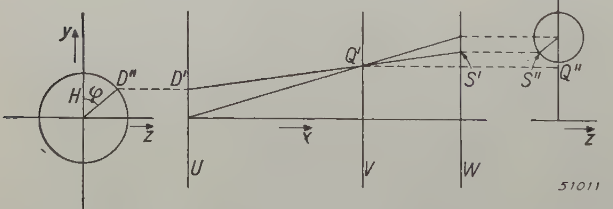


Fig. 2. Construction in the *xy* and *yz* projections of a cone of rays with apex Q (projections Q', Q'') intersecting the plane U in a circle with radius H. This cone intersects the plane W in a circle with respect to which the *yz* projection is displaced from Q to Q'' (adjustment aberration).

We may now call $y - y_0$ and $z - z_0$ $(= z$, since $z_0 = 0)$ the aberrations for the ray determined by y_u, z_u and Q. These aberrations result from the choice of the plane W. In practice they may occur, for instance, with a photographic camera when the distance between the frosted glass plate and the lens is incorrectly adjusted. As a matter of fact the aberrations described by equations (1a, 1b) are sometimes called first-order aberrations, because they are described by an expression of the first degree in H and y_0. There are two kinds of first-order aberrations. The term with y_0 still remains even when the magnitude of H is zero. It then results in an increased distance between the image point and the axis, the increase being in the proportion of $(1 + a_2) : 1$. This may be called the first-order enlargement aberration. The terms governed by H increase with H, thus with the opening of the cone of rays. These may be called first-order aperture aberrations. Obviously both enlargement aberrations and the aperture aberrations are proportional to the distance d between V and W and they disappear when W is coincident with V $(d = 0)$.

Higher-order aberrations

If (*fig. 3*) the point source P is chosen on the axis of the optical system, but the condition that the ray must make a very small angle with the axis is dropped, then owing to the rotation-symmetry the ray will lie entirely in a plane through the axis. After refraction, however, it will no longer intersect the axis in the paraxial image point Q but in a point Q_1 and upon extension meet the plane V at a point S. If, to determine the ray, we again choose a plane U and a point of intersection therein having the coordinates $y_u = H \cos \varphi$, $z_u = H \sin \varphi$, then the coordinates of S will be

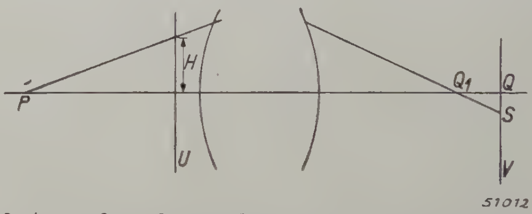


Fig. 3. A ray through P in the plane of the diagram $(\varphi = 0)$, making a large angle with the axis, intersects the axis not at the paraxial image point Q but at Q_1 and the plane V at S, where to a first approximation QS is proportional to H^3. H is the distance from the point of intersection of the ray with the plane U to the axis.

$$y = A \cos \varphi, \quad z = A \sin \varphi, \quad . \quad . \quad . \quad \text{(2)}$$

where A is a function of H, which may be developed in a series of the form

$$A = c_1 H^3 + e_1 H^5 + \ldots \qquad \text{(3)}$$

The series necessarily has odd terms only, because when H is replaced by $-H$ and φ by $\varphi + \pi$, which does not cause y_u and z_u to change in value, y and z must also remain unchanged.

The aberrations (2) are known as **spherical aberration**. The terms with the coefficient c_1 are called third-order spherical aberration, those with the coefficient e_1 fifth-order spherical aberration, and so on.

Let us now consider a point source P off the axis but in the same xy plane (*fig. 4*). If the laws of paraxial imaging were still to hold, the image point would be in Q at a distance y_0 from the axis. An arbitrary ray determined by its point of intersection $y_u = H \cos \varphi$, $z_u = H \sin \varphi$ on a plane U will now strike the paraxial image plane V at a point S, the xy projection (S') and the yz projection (S'') of which are given in fig. 4.

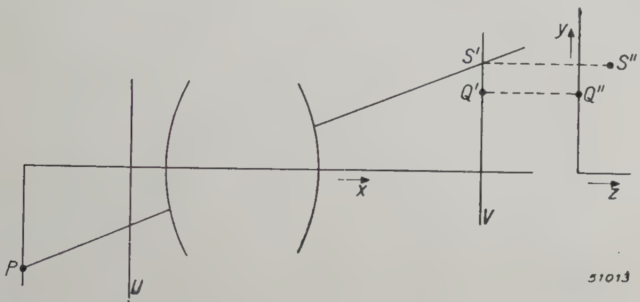


Fig 4. Construction in the xy projection of a ray through P intersecting the plane U at the point $y_u = H \cos \varphi$, $z_u = H \sin \varphi$. P lies in the plane of the figure which at the same time is the plane of symmetry, but the ray in question does not lie in that plane. This ray intersects the plane V at a point S the xy projection S' and the yz projection S'' of which are indicated and which deviates from the paraxial image Q (projections Q' and Q'') of P (see formulae 4a and 4b). The distance from Q' and Q'' to the axis is y_0.

Calculation shows that the coordinates y and z of S are given by

$$y = y_0 + c_1 H^3 \cos \varphi + c_2 H^2 y_0 (2 + \cos 2\varphi) + + c_3 H y_0^2 \cos \varphi + c_4 y_0^3 + \dots \quad (4a)$$

$$z = c_1 H^3 \sin \varphi + c_2 H^2 y_0 \sin 2\varphi + + c_3' H y_0^2 \sin \varphi + \dots \quad (4b)$$

in which the constants $c_1 \dots c_4$ depend not only upon the nature of the optical system but also upon the position of P and the choice of the plane U.

Since we shall here confine our considerations to third-order aberrations the series expansion has been broken off at the terms of the third degree.

If one wishes to find the point of intersection of the same ray on a plane W parallel to V at a distance d, then the right-hand terms of equations (4a) and (4b) are increased by terms of the form (1a) and (1b), whilst if d is small enough the coefficients c may be regarded as remaining unchanged.

The fact that the equations for third-order aberrations are indeed such as given in (4a) and (4b) can most easily be deduced by a reasoning that originated with Conrady [1]).

Let us consider for this purpose (*fig. 5*) refraction at a single convex boundary plane (centre M, radius r, index of refraction left: n_1, right: n_2). We are not only interested, as we were above, in all points of light situated in a plane at right-angles to the axis of the system, but rather in those lying on a sphere with radius R_1 and centre M. Let one of these points be P, so that $MP = R_1$. It is clear that in this special case, owing to the spherical symmetry, the line PM may equally well be considered as an axis as any other. If we confine ourselves to rays which are paraxial with respect to the new axis PM then the image point Q appertaining to it will be at a distance $-R_2$ from M, in which case it follows from the elementary theory that

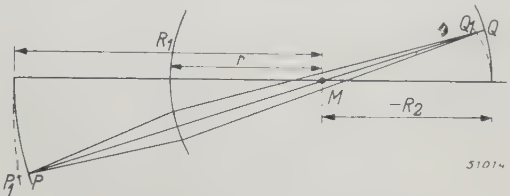


Fig 5. Refraction at a spherical boundary plane (radius r). The point P is imaged by rays paraxial with respect to PM at the point Q. Displacement of P to P_1 causes the image point Q to shift to Q_1.

$$\frac{1}{n_1 R_1} - \frac{1}{n_2 R_2} = \frac{1}{r}\left(\frac{1}{n_1} - \frac{1}{n_2}\right) = -\frac{1}{r}\,\Delta\left(\frac{1}{n}\right).$$

The quantities R_1 and R_2 are to be taken as positive or negative according to whether the concave side of the lens faces to the right or to the left. If P is situated anywhere on the sphere R_1 then the locus of the point Q lies on the sphere R_2.

If we then consider the ray PA (*fig. 6*), which is not paraxial with respect to the axis PM, we find that under the influence of spherical aberration this ray does not cut the sphere at Q but at a point S having as coordinates

$$y' = ka^3 \cos \psi \quad \text{and} \quad z' = ka^3 \sin \psi,$$

in which y' is the distance from the y projection to the axis PM and $z' = z$ is the distance to the plane of delineation, whilst a represents the distance from A to the axis and ψ is the angle between the plane PAM and the plane of delineation.

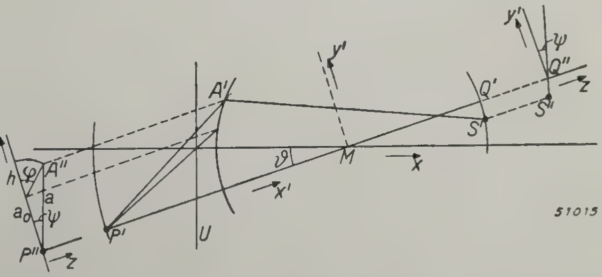


Fig. 6. Explanation of the third-order aberrations arising from spherical aberration according to Conrady. The ray from P to A, of which points the $x'y'$ projections $P'A'$ and the $y'z'$ projections $P''A''$ are indicated, strikes the sphere through Q (projections Q', Q'') at S (projections S', S'').

---

[1]) A. E. Conrady, The five aberrations of lens-systems, Monthly Not. Roy. Astron. Soc. 79, 60-66, 1918.

With the aid of the $\mathfrak{y}'z'$ projection drawn on the left-hand side of fig. 6 we easily find that

$$a \sin \psi = h \sin \varphi \quad \text{and} \quad a \cos \psi = a_0 + h \cos \varphi.$$

The meanings of a_0, h and φ are given in the diagram. Since, further

$$a^2 = a_0^2 + h^2 + 2a_0 h \cos \varphi,$$

we find for S:

$$\mathfrak{y}' = k[a_0^3 + 3a_0^2 h\cos \varphi + a_0 h^2 (1 + 2 \cos^2 \varphi) + h^3 \cos \varphi] \quad (5a)$$

$$z' = k[a_0^2 h\sin \varphi + 2a_0 h^2 \sin \varphi \cos \varphi + h^3 \sin \varphi] \;.\;.\;.\;. \quad (5b)$$

Since it is only a question of small quantities and we break off the expansion at terms of the third degree, the right-hand expressions represent at the same time the aberrations $\mathfrak{y}—\mathfrak{y}_0$ and $z—z_0$ of the point S in the system of coordinates $x\mathfrak{y}z$. (In the transformation from y' to y, one may take $\cos \vartheta = 1$ because higher terms only affect aberrations of the fifth and higher orders.) Bearing in mind, further, that to a sufficient approximation (and again the factors ignored affect only the higher order aberrations)

$$a_0 = \alpha\, y_0 \quad \text{and} \quad h = \beta H,$$

where $\mathfrak{y}_0$, H and φ have the same meaning as before. Then we find

$$\mathfrak{y}—\mathfrak{y}_0 = \gamma_1 H^3 \cos \varphi + \gamma_2 H^2 y_0 (2 + \cos 2\varphi) + \\ + 3\gamma_3 H y_0^2 \cos \varphi + \gamma_4 \mathfrak{y}_0^3, \;.\;.\;. \quad (6a)$$

$$z—z_0 = \gamma_1 H^3 \sin \varphi + \gamma_2 H^2 \mathfrak{y}_0 \sin 2\varphi + \gamma_3 H \mathfrak{y}_0^2 \sin \varphi, \;.\;. \quad (6b)$$

by which equations (4) are obtained, but with this difference that the coefficients of Hy_0^2 in (4) are not, as here, in the ratio 3 : 1. This is because in this case the coordinates are determined from the point of intersection of the refracted ray on the sphere R_2, whilst the image point is chosen on the sphere R_1. If P is moved along MP to P_1 on a sphere with radius R'_1 which is tangent to the sphere with radius R_1, then the point Q is displaced to Q_1. Equations (6) now represent to no less good an approximation (for the term with y_0^3 only is this not absolutely correct) the coordinates on a sphere with radius R_2' tangent to the sphere with radius R_2, for which we again have:

$$\frac{1}{n_1 R_1'} - \frac{1}{n_2 R_2'} = \frac{1}{r}\left(\frac{1}{n_1} - \frac{1}{n_2}\right) = -\frac{1}{r}\,\Delta\left(\frac{1}{n}\right).$$

A flat object plane at right-angles to the principal axis of the system is obtained by taking $R_1' = \infty$. R_2' thereby assumes a certain value R_p, for which the expression

$$-\frac{1}{n_2\, R_p} = -\frac{1}{r}\,\Delta\left(\frac{1}{n}\right)$$

applies.

This is the so-called **Petzval** surface. The aberration in the paraxial image plane, *i.e.* the plane surface at right-angles to the main axis tangent to the sphere R_2, differs from the expression (6) by an amount proportional to the distance from Q_1 to that plane ($= \mathfrak{y}_0^2/2R_p$) and also proportional to the aperture parameter H. When this correction is made an equal amount is added to the two coefficients $3\gamma_3$ and γ_3 and the simple ratio of 3 : 1 is lost.

When dealing with a series of refracting surfaces (a system of lenses) instead of one such surface, one may proceed further and consider the spherical surface with radius R_P as the object plane for the next refraction. One then still arrives at an equation of the form of (6), which then applies to a sphere with radius R_P, for which one finds:

$$\frac{1}{n_{l+1} R_P} = \sum_{i=1}^{i=l} \frac{1}{r_i}\,\Delta_i\left(\frac{1}{n}\right).$$

Here l relates to the last refracting surface; R_P is the radius of the **Petzval** surface of the system. Here again the ratio 3 :1 exists between the coefficients of $H\mathfrak{y}_0^2$ but it is lost again when the corrections are made which are necessary for the transition to a plane image field.

Further considerations of third-order aberrations

A better insight into the significance of the terms of the equations (4a) and (4b) is obtained by studying the intersection figures formed when, with H constant, φ is made to assume all possible values.

It must then be borne in mind that although the analytical treatment proposed gives a complete insight into the geometry of the refracted pencils of rays this is no longer the case if, as we shall do presently, each of the terms is viewed separately. The fact that this has often led to incorrect conclusions being drawn in the literature on the subject has been pointed out already by **Gullstrand** [2]).

Even though one may have an absolutely true picture of the course of the pencils of rays after refraction, one may not immediately draw conclusions therefrom about the distribution of intensity in the beams. Here the approximation of geometrical optics fails more or less, since it does not allow for the coherence between the rays. This has been pointed out by **Picht** [3]) and **Zernike** [4]) among others.

For the sake of simplicity we shall now examine equations (4a) and (4b) one term at a time.

Spherical aberration

In the terms with H^3 one recognises the spherical aberration previously discussed. This therefore remains unaltered for points outside the axis. When this defect occurs alone then as φ varies the point S describes a circle around the point Q, the radius of which is proportional to H^3 (*fig.* 7).

Coma

The terms with H^2y_0 all come under the name of **coma**. If these terms alone were present, then with a variation of φ and constant H, owing to the term $2\,\varphi$, the point of intersection S would describe a circle in such a way that if the zone in the diaphragm plane is passed through once then that circle would be described twice. The centre of this circle lies at a distance $2c_2H^2y_0$ from the paraxial image point. Owing to the proportionality with y_0 one may here speak of an enlargement aberration, which, however, is proportional to H^2. If H be given different values in succession then one obtains (*fig. 8*) a system of circles the common tangents of which intersect each other at an angle of 60°.

---

[2]) See for instance A. **Gullstrand**, Naturwiss. **14**, 653-664, 1926.

[3]) J. **Picht**, Optische Abbildung, Braunschweig, 1931.

[4]) See B. R. A. **Nijboer**, thesis, Groningen, 1942.

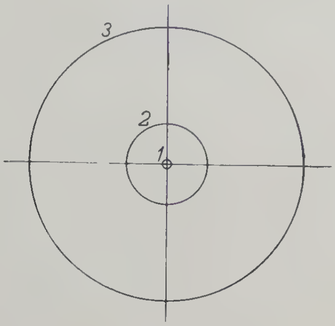


Fig. 7. Intersection figures on the plane V corresponding to the terms with H^3 (spherical aberration) for values of H which are as 1 : 2 : 3.

Astigmatism and field curvature

Owing to the proportionality with H, the terms containing Hy_0^2 are to be regarded as aperture aberrations, which, however, depend upon y_0. The intersection of the rays of a zone (H constant, φ variable) with the plane V is an ellipse with axes $c_3Hy_0^2$ and $c_3'Hy_0^2$. When we consider the intersection upon a plane W at a distance d from V then the terms $a_1H\cos\varphi + a_2y_0$ and $a_1H\sin\varphi$ respectively are added to the aberrations, thereby altering the relation between the axes of the ellipse. By a suitable choice of d one can reduce either the coefficient of $\cos\varphi$ or the coefficient of $\sin\varphi$ to zero, in consesequence of which the figure of intersection converges into a small line at right-angles to the y axis and the z axis respectively (*fig. 9*). The centres M_s and M_t of these degenerate ellipses are called respectively the **sagittal** and the **tangential** or **meridional image points**. The locus of the sagittal image points when y_0 varies is a circle in the xy plane and a convex surface in space, the so-called sagittal image surface; likewise the tangential image points lie on another convex surface, the tangential image surface.

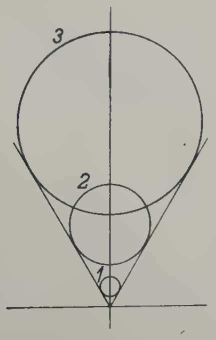


Fig. 8. Intersection figure on the plane V corresponding to the terms with H^2y_0 (coma) for values of H which are as 1 : 2 : 3.

The curvatures of the image surface are given by

$$\frac{1}{R_t} = \frac{2c_3}{a_1/d} \text{ resp. } \frac{1}{R_s} = \frac{2c_3'}{a_1/d} \quad . \; . \; . \; . \quad (5)$$

The equation

$$\frac{1}{R_m} = {}^1/_2\left(\frac{1}{R_s} + \frac{1}{R_t}\right) \quad . \; . \; . \; . \quad (6a)$$

represents what is called the **average field curvature**. The figure of intersection on the sphere corresponding to this curvature is a circle. The quantity

$$A = {}^1/_2\left(\frac{1}{R_t} - \frac{1}{R_s}\right), \quad . \; . \; . \; . \quad (6b)$$

determining the distance of the points M_s and M_t is called the **astigmatism**. An interesting relation arises between R_s and R_t when these quantities are related to the quantity R_P, the so-called **Petzval radius of curvature** of the system.

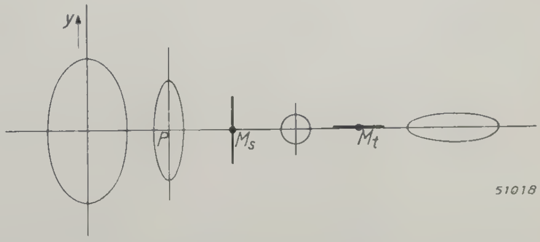


Fig. 9. Series of cross sections (H = constant) with a number of planes W (d negative, increasing in absolute value) corresponding to the terms Hy_0^2 (field curvature and astigmatism). M_s is the sagittal, M_t the tangential or meridional image point. Halfway between M_s and M_t the cross section is a circle. At P (the intersection with the **Petzval**-surface, PM_s = 1/3 PM_t) the ratio of the axes is 1 : 3.

As we have seen above, this is simply related to the data of the system. Instead of the expression given above one may also write

$$\frac{1}{n_{l+1}R_P} = -\sum_{k=1}^{k=l+1}\frac{\Phi_k}{n_k},$$

in which Φ_k represents the strength of the lens (placed in a vacuum) formed by the two successive boundary surfaces, and n_k the index of refraction of the medium between those boundary surfaces. Here Φ_k has to be taken for zero thickness so that

$$\Phi_k = (n_k - 1)\left(\frac{1}{r_{k-1}} - \frac{1}{r_k}\right),$$

whilst $1/r_0$ and $1/r_{l+1}$ have to be taken equal to nil.

The Petzval curvature $1/R_P$ and the two field curvatures are related thus:

$$\frac{1}{R_P}=\frac{1}{R_m}-2A=\frac{1}{R_s}-A=\frac{1}{R_t}-3A\,. \quad (7)$$

Distortion

Finally the term y_0^3 represents an enlargement aberration dependent upon y_0 but independent of H and thus existing for an infinitely narrow beam. It results in a non-linear radial distortion of the image. A square lying in the object plane symmetrically with respect to the axis is distorted into the form of a barrel (barrel distortion) or into the form of a cushion (pincushion distortion) according to whether c_4 is negative or positive.

One must always bear in mind that the aberrations dealt with here and given separate names are equivalent terms in a series expansion and that the true deviation of a point of intersection with respect to the paraxial image point is obtained by summing all the terms. Consequently the true intersection figure of the rays of a zone upon the plane V or upon a plane W may be a complicated curve (in general of the fourth degree). Further it must be remembered that the coefficients c depend not only upon the properties of the system in general but also upon the position of the object and the choice of the diaphragm plane.

Finally it is obvious that the third-order aberrations discussed here represent only an approximation and that to obtain a complete insight it would be necessary to consider also the terms of the fifth and higher orders.

Chromatic aberrations

One would be inclined to deduce from the foregoing that in order to obtain a good image formation in a flat image plane with the desired enlargement it is only necessary to ensure that the system has the right focal distance and that the first and third-order aberrations, *i.e.* the coefficients $a_1 \ldots c_4$, are zero. This would involve satisfying a number of conditions by giving the right values to the curvatures and the distances of the lens surfaces and to the indices of refraction of the kinds of glass, since the coefficients $a_1 \ldots c_4$ are functions of the latter quantities. No allowance would then have been made, however, for the fact that the indices of refraction and hence the coefficients $a_1 \ldots c_4$ depend on the wavelength λ. If the focal distance for a given λ has one value, for some other λ it will be different. Consequently for the same adjustment, thus for the same position of the image plane, even though the adjustment errors for a given λ are nil, for any other wavelength aberrations of the first order will arise. These are called first-order chromatic aberrations. The coefficients $c_1 \ldots c_4$ will also depend upon the wavelength λ. These conditions can be met by requiring that the coefficients $a_1 \ldots c_4$ for a number (two or three) of values of λ must be zero, in the hope that they will not differ much from λ zero for intermediate wavelengths. This, however, considerably increases the number of conditions that have to be satisfied. Moreover, with an objective having a large aperture (large H/f) and an extensive field (large y_0/f) not only the first and third-order aberrations but also those of higher (fifth and seventh) order must be considered.

The designing of a system of lenses

It is evident that from a mathematical analysis of the aberrations alone it is almost impossible to design a system of lenses that has to satisfy high demands. In practice, therefore, one does not proceed in this way at all. The designing of optical systems is more often than not a matter of practical experience rather than one of theory and to some extent more of an art than a science. Nevertheless, the result will always be put to the test by applying an analysis of the aberrations to a solution arrived at empirically. It will then usually be found that the third-order aberrations are not zero but that there is a combination of these together with aberrations of a higher order, such as to make the result on the average as satisfactory as possible.

The difficulties caused by chromatic aberrations do not arise when mirrors are used for the projection. This is one of the reasons why in some cases a mirror system is preferred to one of lenses. Another reason is that the spherical aberration of a concave mirror is less than that of a refracting surface of the same strength.

The spherical mirror

Let us imagine that we have a spherical mirror with circular rim. The axis of symmetry is chosen as the main axis of the optical system. The paraxial rays radiating from a point P_0 on this axis converge upon a point Q_0 on the same axis. A non-paraxial ray from P_0 after reflection meets the axis at a different point Q_0' and reaches the paraxial image plane at a point S, so that Q_0S ($= c_1H^3$) again represents the spherical aberration.

We shall now consider a point source P outside the axis but at the same distance from the centre M as P_0. The line PM for the convex surface of which the mirror forms a part will then likewise be an axis of symmetry. Rays which are paraxial with respect to PM will converge upon PM at a point Q, where $QM = Q_0M$. The rays which are not paraxial with respect to QM strike this line at another point Q'. Thus the aberrations in the imaging of the point P may be described as spherical aberrations with respect to the secondary axis PM.

The situation is now absolutely the same as that in the case described above (see figs, 5 and 6) of refraction by a single convex refracting surface. To this, too, one can apply the theory of Conrady and in this way deduce the values of the coefficients of the third-order aberrations.

Diaphragm at the centre of curvature

We shall not go into the details of the general case but discuss the special case in which the diaphragm plane passes through M. Further we shall choose the distance $P_0M = \infty$ and thus consider the imaging of an infinitely distant plane. The paraxial image point of P_0 then coincides with the focus of the spherical mirror. It is easily seen (*fig. 10*) that owing to the particular position of the diaphragm all beams passing through the diaphragm are subject to the same conditions (apart from the fact that from the point of view of a beam falling obliquely on the circular aperture of the diaphragm the aperture assumes the form of an ellipse having slightly different axes). Rays which are paraxial with respect to PM converge at a distance $|f| = |R/2|$ from M. The other rays have the same spherical aberration with respect to the sphere with radius $|f|$ as that shown by the rays from P_0 with respect to that sphere. We may therefore say that for rays which are paraxial with respect to a secondary axis the image formation is sharp in a curved image plane (radius $|f|$) and that the other rays only display a spherical aberration.

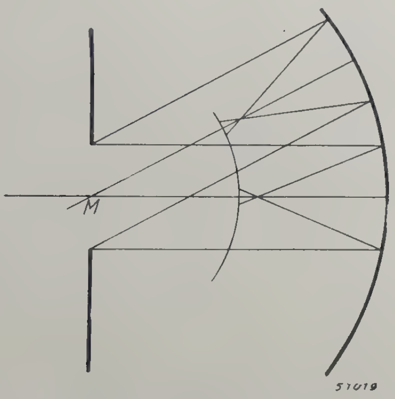


Fig. 10. A diaphragm the centre of which coincides with the centre of the mirror M brings about a practically sharp image on a sphere with radius $|f| = R/2$. Depending on the diameter of the diaphragm, however, spherical aberration occurs. The point sources P_0 and P are to be imagined as the infinitely distant points of the main and secondary axes respectively.

We may, however, also — as was done above for a system of lenses — consider a flat image field through the focus (paraxial image plane with respect to the main axis P_0M) and express the aberration with respect to this plane by the coefficients $c_1 \ldots c_4$. It then appears that when $y_0 = f \tan \vartheta$ ($\vartheta = \angle\, P_0MP$) and $z_0 = 0$ the point of intersection of the reflected ray from P on the image plane is given by

$$y - y_0 = \frac{1}{2R^2} H^3 \cos\varphi - \frac{2}{R^2} H y_0^2 \cos\varphi,$$

$$z - z_0 = \frac{1}{2R^2} H^3 \sin\varphi - \frac{2}{R^2} H y_0^2 \sin\varphi,$$

so that $c_1 = {}^1/_2R^2$, $c_3 = c_3' = -2/R^2$ and $c_2 = c_4 = 0$. From this we find again that although there is field curvature there is no astigmatism nor any coma or distortion. Further the only defect is the spherical aberration.

As may easily be proved, $a_1 = d/f = 2d/R$ and the field curvature is thus

$$\frac{2\,c_3}{a_1/d} = \frac{2}{R}.$$

The radius of the curved image plane is thus $|R/2| = |f|$. Since astigmatism is zero, this surface is at the same time the Petzval surface. This may be deduced, according to a common method from the general formulae given above for the Petzval curvature, which apply for the case of refraction, by putting n prior to reflection $= 1$ and after reflection $= -1$, so that $\Delta\,(1/n) = 2$.

Spherical mirror with correction plate

If when using a spherical mirror one had a means of causing the incident rays to change direction when passing through the diaphragm plane in such a way as to eliminate the spherical aberration, then a perfect projection would be obtained on a sphere with radius $|R/2|$.

Such a possibility is of great practical importance. The means of attaining this to a high degree of approximation is to be found in Schmidt's correction plate, which is optically flat on one side and has an aspherical surface on the other side.

It is clear that with this a system can be devised which combines the advantages of a large aperture with those of a wide image field. This provides a simple way of obtaining a result which would otherwise entail a highly complicated system of lenses. It is even possible to exceed considerably the light flux of the present lens systems. One must, however, remember that the image field is not flat but curved.

The construction and applications of the correction plate will be dealt with in subsequent articles.

Appendix: Note on the parabolic mirror

Finally some remarks are to be added regarding the parabolic mirror.

Fig. 11 represents the cross section of the plane of the diagram for a parabolic mirror whose radius of curvature at the top is R. If the normal is drawn from a point P on the mirror surface it will intersect the axis at a point N, such that the distance from N to the foot Q of the perpendicular from P on the axis (the so-called sub-normal) is R, contrary to the case with the sphere with centre M, where the length of the normal itself is R while all normals pass through M. This difference in behaviour of the two curves has the effect that rays striking the axis obliquely show no spherical aberration but all pass exactly through the focus F, whereby $MF = FO = {}^1/_2\ |R|$. Without further thought one might suppose that a parabolic mirror with a diaphragm at the centre of curvature M would automatically be free from aberrations, but this is not so, because with the parabola there is not such a high degree of symmetry with respect to the point M as there is with the sphere. In fact one finds by calculation that

$$c_1 = 0,\ |c_2| = 1/R^2,\ |c_3| = 8/R^2,\ |c_3'| = 4/R^2,\ |c_4| = 4/R^2.$$

Except for spherical aberration — which is naturally absent — all other aberrations do indeed occur.

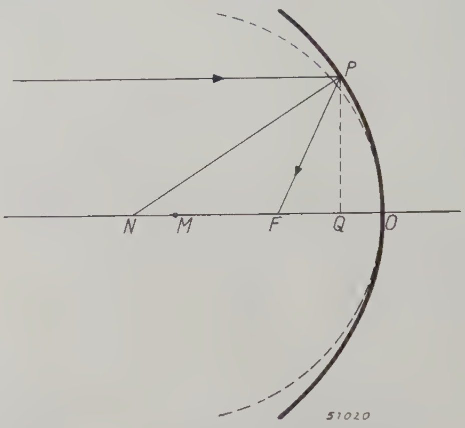


Fig. 11. The passage of rays with a parabolic mirror. A ray through P parallel to the axis, after reflection, strikes the axis at F, $MF = FO = {}^1/_2\ |R|$. Here R is the radius of curvature at O (broken line). NP is the normal drawn from P on the parabola; $NQ = MO = |R|$.

CONSTANT AMPLIFICATION IN SPITE OF CHANGEABILITY OF THE CIRCUIT ELEMENTS

by J. J. ZAALBERG van ZELST.

621.394.645.3

In a previous article [1]) amplifier circuits were discussed which make the amplification very little dependent upon the slope of the valves, assuming the other circuit elements to be constant. The present article deals with the problem of designing such a circuit as to minimise also the effects of variations in those other circuit elements upon the amplification, or, to be more precise, to keep the amplification constant within limits narrower than the tolerances of the variable elements. The solution of this problem is built up upon the method described in the previous article, where the amplifier is caused to oscillate with a non-disturbing auxiliary frequency. As feedback network a quadripolar system is used which in the event of variations taking place in the component parts possesses a practically constant attenuation, provided that at the same time the (auxiliary) frequency is modified in such a way that the phase difference between the input and output voltages of the quadripolar system remains constant. This condition is automatically satisfied in the oscillator, since the latter can only oscillate when the said phase difference amounts to 180°. In this manner circuits can be made with an amplification per valve of, for instance, approximately 8 ± 1% for a tolerance of + or — 5% in the values of the component parts. A number of such stages can quite well be placed in cascade.

Introduction

An amplifier consists of one or more amplifying valves and a number of other circuit elements such as resistors, capacitors, etc. The magnitude of the amplification depends, as far as the valves are concerned, mainly upon the slope (mutual conductance) and, as regards the other elements in a given circuit, upon their impedance. In a previous article in this journal [1]), which we shall further refer to as article I, we discussed a number of circuits where variations in slope (within certain limits) result in a very much smaller relative change in amplification. For all practical purposes we were therefore able to speak of "constant" amplification. It was expressly presupposed, however, that the values of the other circuit elements were constant.

We will now go a step further and aim at altering at least one of the circuits discussed in such a manner as to make the amplification very little subject to the influence of the values of the resistors, capacitors, etc., while retaining of course insensitivity to changes of the slope.

Let us first for a moment consider another problem, the construction of a direct voltage source of which the voltage remains constant within very narrow limits. To arrive at a solution at all practicable we should have to have recourse to some material constant, for instance an electro-chemical EMF (in the case of a standard cell) or a normal cathode fall (in the case of a stabilizing tube). Now in practice it is impossible to get conditions under which material constants are actually constant, for as a rule they are dependent upon other physical or chemical quantities over which one has not complete control, e.g. the temperature, or the purity of the materials used, which are apt to change in the course of time and have their effect upon the result. But even if the principle of the voltage source is based upon suitably chosen material constants, the inconstancy of the voltage cannot be made smaller than that of the practically least variable material constant — both inconstancies being taken, of course, relatively.

What has been said here not only applies to voltages but likewise holds for many other physical quantities. One might, therefore, easily be led to ascribe general validity to it. In the case of the amplifier already mentioned above one would also be inclined to assume *a priori* that at a given frequency the amplification cannot be more accurately determined than the slope, resistance, capacitance, etc. deviating the least from the prescribed value (with the possibility of all these quantities varying). Amplification, however, is a non-dimensional quantity; it may well be found less difficult to imagine a relation of two similar quantities as having a constant value than to suppose the same of a physical quantity itself, but it is not so simple to be able to say why the former is more readily acceptable. However this may be, it does indeed appear possible to construct alternating voltage amplifiers whose amplification satisfies the requirement of stability, that is to say whose amplification differs from a given value relatively less than the relative deviation that we wish to allow for each of the component parts from their prescribed values. In other words, if in such an amplifier a tolerance of + or — $p\%$ is allowed in the values of the component parts, then the amplification obtained should deviate from the nominal value by less than $p\%$. How this is to be attained is the subject of this article.

---

[1]) J. J. Zaalberg van Zelst, Stabilised Amplifiers, Philips Techn. Rev. 9, 25-32, 1947 (No. 1).

Here we will proceed to build up upon the method discussed sub IIb in article I, where the amplifier oscillates with a non-disturbing auxiliary frequency.

The amplifier with negative feedback

In addition to this method there is another means by which it seems at first sight possible to arrive at the object in view. We have in mind an amplifier possessing a feedback resistance R_k in the cathode line and a resistance R_a in the anode circuit, where the voltage on these two resistances in series serves as the output voltage V_0. By choosing SR_k large in comparison with unity ("infinitely strong feedback"; S = slope of the valve) the amplification becomes approximately equal to $1 + R_a/R_k$ and thus independent of S. Moreover, if R_a is chosen only small with respect to R_k, it can be made as little dependent upon the resistance values R_a and R_k as one wishes; the amplification then drops, however, to scarcely more than unity. If we consider how large R_a may be taken while still satisfying the requirement of constancy, we find $R_a = R_k$; the nominal amplification is then 2, thus still only very modest. Although amplifiers with such a low amplifying factor may be useful in certain cases, we shall leave them out of consideration here partly for the following reason.

If one tries to reach higher amplifications by employing stages in cascade one finds that the amplification still does not work out higher than a nominal value of 2 so long as the constancy requirement is complied with — apart from the practical difficulties connected with a cascade circuit of stages not having a common input and output terminal.

The oscillating amplifiers

We would recall — see for instance article I — that in order to make a valve oscillate a feedback is necessary, bringing a voltage derived from the anode current to the input of the circuit. In *fig. 1* it is indicated diagrammatically how a voltage V_{t1} when brought to the input of an amplifier A gives rise to an anode current I_a from which the feedback network T derives a voltage V_{t2}. When the output terminals of T are properly connected with the input terminals of the amplifier the condition will be maintained (*i.e.* the amplifier begins to oscillate) if $-V_{t2} = V_{t1}$. If we define the (total) slope of the amplifier as $S = I_a/V_{t1}$ and the transfer impedance Z of the feedback network as $Z = V_{t2}/I_a$, then $V_{t2} = ZI_a = ZSV_{t1}$. The condition that must be satisfied for $-V_{t2} = V_{t1}$ is thus:

$$ZS = -1 \quad \ldots \ldots \ldots \quad (1)$$

The minus sign expresses that a phase difference of 180° must exist between V_{t1} and V_{t2} (where positive directions are chosen as indicated in fig. 1).

In article I we discussed circuits where the relation (1) is used to determine the slope S of the amplifier as the reciprocal of a given transfer impedance Z in spite of possible changes in the valve slope. This was found to be attainable in the best way by causing the valve to oscillate with a non-disturbing auxiliary frequency. From the oscillating voltage a regulating voltage was derived which compensates the changes in slope.

As a rule the transfer impedance Z is a function of the frequency. For the frequency at which relation (1) is satisfied, *i.e.* for the oscillation frequency, the rule therefore applies that the total amplification in the circuit formed by the amplifier and the feedback network in cascade is exactly equal to unity; in other words the amplification in the amplifier itself is just compensated by the attenuation in the feedback network.

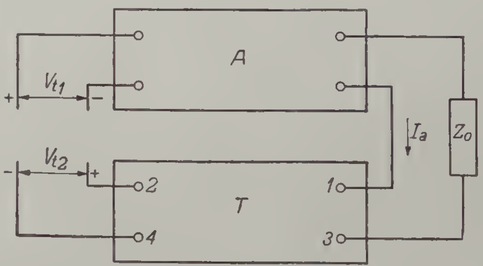


Fig. 1. If for a certain frequency the transfer impedance $Z = V_{t2}/I_a$ of the feedback network T assumes the value $-1/S$ (S = slope of the amplifier A, $= I_a/V_{t1}$) then the voltage supplied by T is just identical with the voltage required at the input of the amplifier to produce V_{t2}. When the output of T is connected to the input of A the amplifier starts oscillating in that frequency.

If one succeeds in stabilizing this attenuation in spite of changes in the circuit elements — and it will be seen that this is possible owing to a property of certain networks — then one will at the same time have obtained a constant amplification.

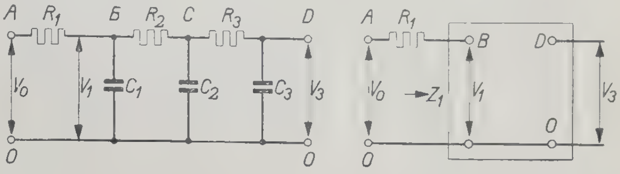
Correlation between the amplitude and phase characteristics of some quadripolar systems

By the amplitude characteristic of a quadripolar system we understand the absolute value of the ratio of the output voltage to the input voltage as a function of the frequency, and by the phase characteristic is understood the phase difference φ between these two voltages likewise as a function of the frequency. If one or more elements in a quadripolar system are altered then at a given frequency f_0 the phase angle φ will as a rule change. By giving the frequency a certain value f' the phase angle φ can be returned to its original value. The correlation just referred to which occurs with certain quadripolar systems consists in the fact that *at the frequency f' the attenuation of the quadripolar system with the altered components is practically equal to that of the quadripolar system with the original values at the frequency f_0*. We shall presently give an example of a quadripolar system possessing this property.

If we now use such a quadripolar system as feedback network so arranged that the phase difference φ becomes 180° for a certain frequency (f_0) then the circuit will oscillate with that frequency (assuming that the amplification is adequate). If we then change one or more elements of the quadripolar system the oscillation frequency will automatically assume the value — just called f' — at which φ remains constant (in this case 180°, since this is one of the conditions for oscillation). Thus we have obtained a circuit which in the event of a change in the elements varies its oscillating frequency in such a way as to keep the attenuation of the feedback network practically constant. *A fortiori* this will therefore also be the case with the amplification, since this, according to another condition for oscillation, is just as great as the attenuation.

If one should now succeed in conducting the signal to be amplified — whose frequency strongly deviates from the oscillation frequency — to the amplifier and carrying off the amplified signal in such a way as to retain the above-mentioned property of the feedback network, then one has indeed reached an amplification substantially constant for the signal frequency.

Before going further into these signal circuits we shall give an example of a quadripolar system possessing the property described above; see *fig. 2a*. The fact that it does indeed possess this property can be proved by a not difficult but extensive calculation. Rather than follow this calculation here we shall demonstrate how the said correlation between the amplitude and phase characteristics comes about with the circuit according to fig. 2*a*. Here we assume that R_1C_1, R_2C_2 and R_3C_3 have been so chosen that for a certain frequency $f_0 = \omega_0/2\pi$ there arises in each of the three sections a phase shift of 60° between the output and input voltages of a respective section. Further it is assumed that $R_3 \gg R_2 \gg R_1$, so that the load of one section upon the previous one may be ignored. Under these simplified conditions the vector diagram of *fig. 3* applies. Since the voltage vectors OB and BA are at right-angles to each other, B lies on a circle with OA as diameter. Similarly one finds that C lies on a circle with OB as diameter and D on a circle with OC as diameter. Owing to the phase shift of 3 × 60° OD comes to lie in the extension of AO; in other words the output voltage V_3 is in counterphase to the input voltage V_0, so that the circuit in which this quadripolar system acts as feedback network will oscillate with the frequency in question (the frequency f_0 for which the phase shift per section of the network is 60°, thus for which $2\pi f_0 C_1 R_1 = \tan 60°$ applies).



a *b*

Fig. 2. *a*) Example of a network possessing the following property: when one or more elements change then the ratio V_0/V_3 of the input voltage to the output voltage does not, in the first instance, change provided the frequency is at the same time changed in such a way that the phase difference between V_0 and V_3 returns to its original value.

b) The same network as in (*a*) but with the part $OBOD$ drawn as a separate quadripolar system, the impedance of which, viewed from OB, is called Z_1.

If, now, one or more of the circuit elements is changed then the oscillating condition $\angle AOD = 180°$ remains valid. The oscillation frequency will therefore adjust itself in such a way that in spite of the alteration of the elements the sum of the angles AOB, BOC and COD is still 180°. The angles

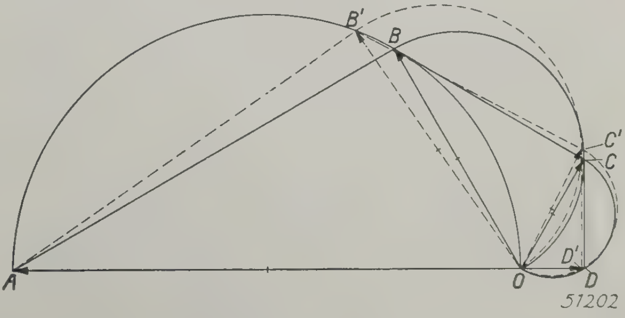


Fig. 3. Vector diagram of the voltages in the circuit of fig. 2*a* when each of the sections R_1-C_1, R_2-C_2 and R_3-C_3 produces a phase shift of 60° between the input voltage and the output voltage of the respective section (corresponding to the angles AOB, BOC and COD) and $R_3 \gg R_2 \gg R_1$. When the elements are changed and at the same time also the frequency is changed in such a way that OD continues to lie in the extension of AO, a figure is produced like that shown in broken lines, where B moves along the circle to B', C to C' and D to D'. This broken-line figure has been constructed for the case where R_1 becomes 25% smaller while the other elements have their nominal values. OD' appears to be only 2% smaller than OD.

may, however, deviate slightly from 60°, B moving along the circle described on AO, for instance to B'. Although OB' may differ noticeably from OB in length the projection from OB' towards a direction OC' approximately coinciding with that from OC will be practically equal to OC, because BC is a tangent to the circle. This applies to a still greater degree for the final result, the projection OD' from OC' on the extension of AO, which then also coincides by good approximation with the original vector OD.

As may easily be seen from fig. 3, the nominal attenuation of the network, *i.e.* the absolute value of the ratio of the input voltage to the output voltage, is equal to $1/\cos^3 60° = 8$.

Example

Applying the index 0 to indicate the values appertaining to the condition in which all elements of the circuit have their nominal value, then:

$$C_{10}R_{10} = C_{20}R_{20} = C_{30}R_{30}, \text{ put} = T.$$

When the elements are varied $C_{10}R_{10}$ changes into $C_{10}'R_{10}'$, say $= p_1T$, $C_{20}R_{20}$ into p_2T, and so on. As a consequence the angular frequency ω_0 of the oscillation will assume a value ω', say $q\omega_0$, and the angles $AOB = \alpha_0$, $BOC = \omega_0$ and $COD = \gamma_0$, originally 60° each, become α', β' and γ'. Also these last-mentioned angles sum up to 180°:

$$\alpha' + \beta' + \gamma' = \pi. \quad . \; . \; . \; . \; . \; . \; . \; (2)$$

Now $\tan\alpha' = \omega C'_1R_1' = q\omega_0 \cdot p_1T = p_1q\omega_0T$, $\tan\beta' = p_2q\omega_0T$, $\tan\gamma' = p_3q\omega T$, whilst $\omega_0T = \omega_0C_{10}R_{10} = \tan\alpha_0 = \sqrt{3}$. Equation (2) thus becomes:

$$\tan^{-1}(p_1q\sqrt{3}) + \tan^{-1}(p_2q\sqrt{3}) + \tan^{-1}(p_3q\sqrt{3}) = \pi. \quad (3)$$

For given variations of the elements, thus for given values of p_1, p_2 and p_3, we get from (3) the value of the factor q by which the frequency changes. With this value given one can calculate the angles α', β' and γ', the cosines of which determine the magnitude of the attenuation; for the ratio of the input voltage to the output voltage of the network (fig. 3) is:

$$OA/OD' = 1/\cos\alpha' \cdot \cos\beta' \cdot \cos\gamma' \; . \; . \; . \; . \; . \quad (4)$$

By way of example let us suppose that only R_1 changes and that this resistance becomes 25% smaller. Then $p_1 = 0.75$, $p_2 = p_3 = 1$. From (3) we then find $q = 1.106$, from which it follows that $\alpha' = \tan^{-1}(0.75 \cdot 1.106 \cdot \sqrt{3}) = 55°9'$, $\beta' = \gamma' = \tan^{-1}(1.106 \cdot \sqrt{3}) = 62°\ 25.5'$. According to (4) the attenuation is now 8.167, *i.e.* only 2% greater than the original value of 8.000, in spite of the great change (25%) in R_1.

Choice of the angles α_0, β_0 and γ_0

Let it be asked that the nominal values of the angles α, β and γ shall be so chosen that when these angles — which together always make 180° — vary the product $\cos\alpha \cdot \cos\beta \cdot \cos\gamma$ shall change as little as possible. To that end the nominal values (α_0, β_0, γ_0) will have to be each 60°, as is to be understood from the following.

Let us first suppose that γ = constant, $= \gamma_0$. Then

$$\cos\alpha \cdot \cos\beta \cdot \cos\gamma = \cos\alpha \cdot \cos(\pi - \alpha - \gamma_0) \cdot \cos\gamma_0 =$$
$$= \frac{-\cos(2\alpha + \gamma_0) + \cos\gamma_0}{2} \cdot \cos\gamma_0.$$

In this equation the only term containing the variable angle α is $\cos(2\alpha + \gamma_0)$. This is the least affected by changes in α when $2\alpha + \gamma_0 = 0$ or π, from which it follows that $\alpha = {}^1/_2\,(\pi - \gamma_0)$ so that $\beta = \pi - \alpha - \gamma = {}^1/_2\,(\pi - \gamma_0)$, thus $= \alpha$.

Following an analogous reasoning it can be proved that if γ is also taken to be variable then the nominal value $\gamma_0 = \alpha$, thus that each of the angles must be chosen at 60°.

In *fig. 4* we have as feedback network a quadripolar system very much resembling that of fig. 2*a*. There is, however, a small difference: as will be seen, in fig. 4 the resistor R_1 is connected parallel to C_1, which is not the case in fig. 2*a*. If we introduce an impedance Z_1 as indicated in fig. 2*b* then:

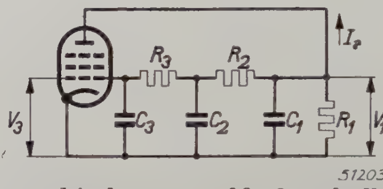


Fig. 4. The quadripolar system of fig. 2*a* as feedback network in a slightly modified form (R_1 parallel to C_1).

$$V_1 = \frac{Z_1}{R_1 + Z_1}\, V_0.$$

In the scheme of fig. 4 the anode current I_a flows through the parallel circuiting of the resistor R_1 and the impedance Z_1, so that here the voltage V_1 amounts to

$$V_1 = -\frac{R_1Z_1}{R_1 + Z_1}\, I_a.$$

Thus the two circuits are equivalent provided $V_0 = -R_1I_a$. From fig. 4 it is also seen that $I_a = SV_3$, and from fig. 3 that $V_0 = -V_3/\cos^3 60° = -8\,V_3$. The condition $V_0 = -R_1I_a$ therefore leads to

$$-8V_3 = -R_1SV_3,$$

or

$$R_1S = 8. \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (5)$$

From the oscillation condition $ZS = -1$ it follows that in combination with (5) the transfer impedance $Z = -R_1/8$, as may easily be deduced direct from fig. 4. In the first instance R_1 is therefore the only element in the network that governs Z. This is another way of formulating the property of the network referred to above.

In order to satisfy this condition also when the valve slope or the resistance R_1 changes, a regulating voltage derived by rectification from the oscillating voltage is caused to act upon one of the grids of the valve. The circuit then oscillates with an amplitude such that the regulating voltage corresponds to the required slope. By taking suitable measures, which we shall not enter upon here ²), it can be ensured that this will be the case already at a very small amplitude of oscillation.

As a consequence the amplified signal — about which more will be said under the next heading — will not be modulated to any extent by the oscillation voltage.

Applying an input circuit and an output circuit

We shall now use the relation $R_1S = 8$ to amplify the signal eight times with the oscillating valve. The frequency with which the amplifier oscillates must be so chosen — as already remarked above — as to prevent any troublesome interference arising, thus for instance much lower than the signal frequency ³).

Let us first consider the circuit represented in *fig. 5*, which has been obtained from that of fig. 4 by adding the circuits L_1—K_1 and L_2—K_2 tuned to the signal frequency. The signal voltage V_i laid on to the input gives rise to an anode current $I_a' = SV_o$ (I_a is used to distinguish this from the anode current I_a having the oscillation frequency). The output voltage V_o then becomes $V_o = Z_2SV_i$, where Z_2 represents the impedance of the circuit L_2K_2 for the signal frequency. Then, however, the signal amplification $V_o/V_i = Z_2S$ is proportional to Z_2 and thus, however constant S may be kept, is dependent upon the elements of the circuit L_2—K_2.

Moreover, with the circuit according to fig. 5 neither of the signal circuits is earthed, and this constitutes an objection if it is desired to connect several stages in cascade. (One could, it is true, couple the output of one stage, for instance, inductively with the input of the next stage, but then we should have the coupling coefficient acting as a factor the variations of which are not compensated.)

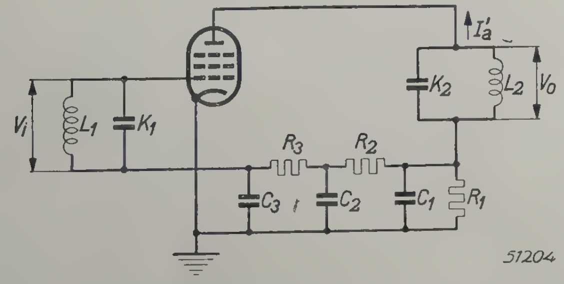


Fig. 5. The same circuit of fig. 4 with the addition of two circuits L_1-K_1 and L_2-K_2 tuned to the signal frequency.

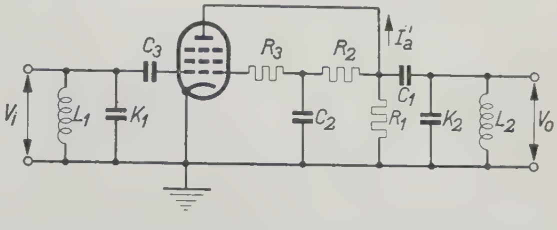


Fig. 6. The circuits L_1-K_1 and L_2-K_2 connected in series with C_3 and C_1 respectively in such a way that both are earthed and the resistor R_1 is connected parallel to the circuit L_2-K_2.

These objections can be overcome by proceeding according to the diagram of *fig. 6*, where the circuit L_1—K_1 is connected in series with C_3 and L_2—K_2 in series with C_1 in such a way that both circuits are earthed. If, now, the oscillation frequency is much lower than the signal frequency then the self-inductances L_1 and L_2 will practically act as shortcircuits for the currents with the oscillation frequency, so that as regards the latter the feedback network may be considered as unchanged. The signal current I_a' now traverses an impedance formed by the parallel circuiting of R_1, R_2, L_2 and K_2 (C_2 and C_1 may be imagined as being short-circuited for the signal frequency). Since the circuit L_2—K_2 behaves, in respect to the signal frequency to which it is tuned, as a resistance which is much greater than R_1, and since, according to a supposition already previously made, R_2 is also large in comparison with R_1, the anode impedance for the signal frequency equals approximately R_1. Hence the output voltage is by approximation $R_1I_a' = R_1SV_i = 8V_i$, so that the amplification does indeed amount to approximately 8 and is practically independent of the values of all the circuit elements employed.

It is to be remarked here that the amplification is not limited just to one value, *i.e.* 8. This is the value found from $1/\cos^3 60°$. In the first place it is not necessary to obtain the phase shift of 180° in three sections of 60° each, for this can also be done with four sections of 45°, five of 36°, etc. or in general with n sections of $180°/n$. The amplification then becomes $1/\cos^n (180°/n)$, *i.e.* for $n = 3, 4, 5, \ldots$ respectively 8, 4, 2.89, . . .; thus it is greatest for $n = 3$, in which case the circuit is also the simplest.

Other possibilities are presented when it is borne in mind that oscillation can be brought about not only with a total phase shift of 180°, but likewise with an odd multiple of 180° and, if an oscillator of two valves is employed, also with even multiples of 180°. For instance in the case of two stages a phase

---

[2] What is meant is that the oscillating voltage can be amplified with an auxiliary valve before a regulating voltage is derived from it by means of rectification; see fig. 7 of article I.

[3] One can also work with an oscillating frequency much higher than the signal frequency provided the necessary alterations are made in the circuit.

shift of 360° can be obtained with 5 sections of 72°, which would lead to an amplification of $1/\cos^5 72° \approx 355$.

Finally it is to be remembered that in the foregoing circuit diagrams we have ignored the possibility of a mutual load between the *RC* sections, it being supposed that $R_2 \gg R_1$, $R_3 \gg R_2$ and so on. If we drop this restriction, which has only been introduced for the sake of simplicity, then the amplification may deviate rather considerably from the above mentioned value $1/\cos^n (m \cdot 180°/n)$, in which m represents the multiple of 180° used. In our example with 3 × 60° one is not restricted to an 8-fold amplification but can give it any value between about 4 and 10 by choosing suitable values.

A practical example

Thus we have indicated in broad lines the fundamental idea of the process to be followed, which may be summarized as follows: just as in article I, an amplifier is employed which oscillates with a non-disturbing auxiliary frequency, while the oscillation voltage supplies a regulating voltage that compensates variations in slope. As feedback network we have a quadripolar system possessing the above-mentioned correlation between the amplitude and the phase characteristics, so that when changes take place in the values of the components the attenuation in this quadripolar system remains practically constant. Equal to this attenuation is the amplification of the signal.

In actual fact account must of course be taken of the factors which have so far been ignored, for which small corrections have to be made. For instance allowance will have to be made for the loads of the second and third *RC* sections (see fig. 2*a*) on the preceding section(s). Account must also be taken of the influence, be it ever so small, of the series resistance of the coils upon the oscillator, and also of the variation in the high-frequency anode impedance due to the deviations allowed in the values of the circuit elements. It would lead us too far to go into all these finesses here.

Suffice it to give here an example of practical application and to mention the maximum deviation to be expected from the nominal amplification. This example is by no means claimed to be the most favourable case that could be found.

In *fig. 7* we have a diagram of an amplifying stage for an angular frequency of $3 \cdot 10^6$ rad/sec. This diagram is complete except for the regulating voltage mechanism, for which an auxiliary amplifier can be used (*cf.* footnote [2])). The values of the components given in the text underneath the illustration are to be regarded as prescribed values. We shall assume that in the actual construction components may be used the value of which is not greater than 1.05 times and not smaller than 1/1.05 times the prescribed value; thus, for instance, where a capacity of 1000 pF is indicated a capacitor of between 1050 and 952.4 pF may be used. Under these conditions the amplification of this stage will deviate less than 1% plus or minus from the nominal amplification, which amounts to 8.1336 [4]).

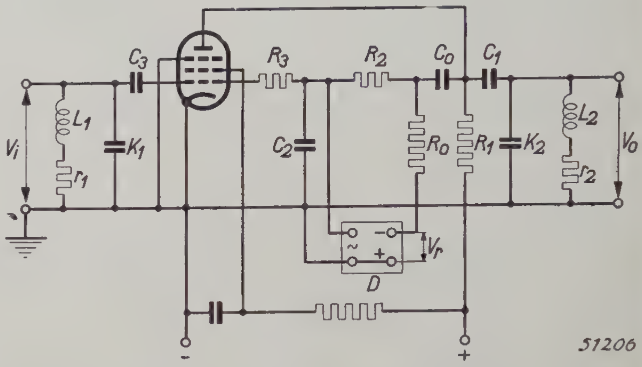


Fig. 7. The circuit of fig. 6 in a further developed stage. From the oscillation voltage on C_2 the part D (not shown in detail) derives a regulating voltage V_r which is brought, *via* the resistors R_0, R_2 and R_3, to the control grid so as to limit the amplitude of oscillation and compensate the slope variations. C_0 is a blocking condenser. The resistance of the coils L_1 and L_2 is indicated by r_1 and r_2 respectively; the other symbols are the same as in fig. 6.

As an example the following case is taken. If the values of the components lie between 1.05 times and 1/1.05 times the prescribed values, then with a signal frequency of $3 \cdot 10^6/2\pi$ the amplification is 8.1336 ± 1%, for which the following components are required:

R_0 = 1 Mohm
R_1 = 2 004 ohm
R_2 = 30 000 ohm
R_3 = 750 000 ohm
C_0 = 10 μF
C_1 = 394 000 pF
C_2 = 24 000 pF
C_3 = 1 000 pF
$L_1 = L_2$ = 2 mH
$r_1 = r_2$ = 30 ohm
$K_1 = K_2$ = 55,5 pF
Nominal valve slope without regulating voltage S > 6 mA/V
Internal resistance of the valve R_i > 1 Mohm
The angular frequency of the oscillation = abt. 2300 rad/sec.

Stages in cascade

From the fact that the amplification per stage varies so much less than the variations of the component parts it follows that a certain number of stages can be connected in cascade without fear of the deviation in the total amplification exceeding the tolerance of the elements. Consequently in the practical example just given, where the ratio of these relative deviations is more than 5, one may go as far as five stages in cascade. One then reaches an amplification of well over 35000 times, exact to within a tolerance of less than 5%

---

[4]) At first sight there does not seem to be any sense in giving the amplification figure to four decimals when deviations up to 1% are possible. It would, however, be incorrect to round off say 8.1336 to 8.13, because the limit values of 8.13 ± 1% would not correspond to the guaranteed limit values of 8.1336 ± 1% (or rather: 8.1336 ± 1.00%).

in the least favourable case. If the values chosen for the elements are well within the tolerance then the deviation from the nominal amplification may be appreciably less.

With these stages in cascade the output circuit of one stage can serve as input circuit of the next stage, thanks to the fact that all these circuits are earthed.

The oscillation frequencies should be chosen far enough apart as to ensure that the stages do not interfere with each other.

What is the least favourable case?

One might well expect the least favourable case to be that where all the circuit elements show the maximum permissible deviation, perhaps with some in the positive direction and others in the negative sense, but this need not always be the case.

In order to demonstrate this in a simple way we will suppose for a moment that only two of all the circuit elements, for instance one resistor R and one capacitor C, have a certain tolerance and that all the other parts have the prescribed value. With R and C as coordinates we can draw lines of constant amplification, which, as the calculation proves, are curves as shown in *fig. 8*. From this diagram it is to be seen that the maximum amplification, in this case 8.3, is reached when R has the minimum value R_{min} and C the maximum value C_{max}, whereas the minimum amplification, 8.0, is obtained with $R = R_{\text{max}}$ and $C =$ a value between C_{min} and C_{max}; with $R = R_{\text{max}}$ and choosing $C = C_{\text{min}}$ or C_{max} the amplification was increased again (up to about 8.15 and 8.05 respectively).

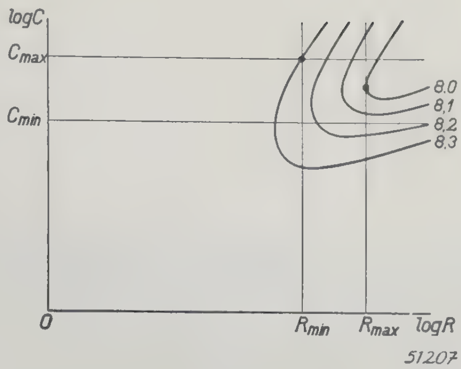


Fig. 8. R_{min}, R_{max}, C_{min}, C_{max} = limits between which the value of one resistor R or of one capacitor C respectively may lie in the diagram of fig. 7; the other components have the prescribed values. With log R and log C as coordinates the lines of constant amplification are the curves indicated. It is seen that the smallest amplification (8.0) is obtained when C has a value between C_{min} and C_{max}.

In principle the same applies when there are more than two variable elements, only then we have to deal with just as many coordinates as there are variable elements.

The Eindhoven Observatory, which for a number of years already has been cooperating with the scientists attached to the Philips Physical Laboratory.

MEASURING THE RATE OF WATCHES WITH A CATHODE-RAY OSCILLOGRAPH

by H. van SUCHTELEN.

681.112.47.001.4: 621.317.361

A description is given of an experimental apparatus with which it is possible to compare the rate of a timepiece with a standard frequency with the aid of a normal cathode-ray oscillograph. By means of a microphone the ticking of the timepiece is converted into voltage peaks which produce the vertical deflection on the oscillograph. The time-base voltage of the oscillograph is synchronized with a fixed standard frequency (60 c/s) derived by frequency division from the highly constant frequency of a quartz oscillator. From the speed at which the peak moves across the oscillograph screen it can be determined how much the timepiece is gaining or losing. A small rate error, of for instance a few seconds per 24 hours, can be determined within a very short space of time (a few minutes).

The usual practice is to test the rate of a watch by comparing it with that of a standard chronometer. To reach an accuracy of say 1 second per 24 hours (for which the watch must of course have a seconds hand) it is necessary to keep the timepiece under observation for 24 hours. After adjustments have been made it is as a rule necessary to check the rate again. Furthermore it is often desired to ascertain to what extent the rate is affected by the position of the timepiece and by the fact whether the spring is wound up or run down. A thorough test may therefore take several days and in many cases this obviously has its objections.

A much quicker method is possible by electrical means, by comparing the frequency of the escapement with a considerably higher standard frequency. This method is applicable not only to complete watches but also to the escapements alone, which are generally made in special factories and have to be tested before delivery to the watchmakers. Apparatus with which these tests can be made have been known for a number of years already. In the Philips laboratory an experimental apparatus has been devised for this purpose which makes use of a cathode-ray oscillograph and which is briefly described in this article.

The watch or escapement 1) to be tested is placed in a holder and clamped onto a microphone which picks up the sound of the ticking. This signal, consisting of a strongly damped vibration and some background noise, is amplified about 10 000 times and applied to a relay valve (a gas-filled triode; see *fig. 1*), in the anode circuit of which sharply defined current impulses are released corresponding to the ticking, which normally should have one of the standardized frequencies of 4, 5 or 6 c/s.

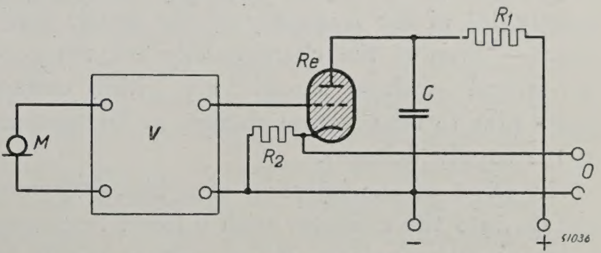


Fig. 1. A circuit supplying sharply defined voltage impulses with the frequency of the ticking of the watch being tested. The circuit formed by the resistor R_1, the capacitor C and the relay valve *Re* connected in parallel to the latter can generate a relaxation oscillation with a frequency approximately equal to the ticking frequency of the watch (4-6 per second). *Via* the microphone M and the amplifier V the ticking produces an alternating voltage on the grid of the relay valve, which synchronises the relaxation oscillation with the ticking. At every tick contact is made in the relay valve and the capacitor C is discharged across the valve and the resistor R_2. The capacitor is then recharged *via* the resistor R_1, discharged again upon the following tick and so on. In this way R_2 receives voltage impulses which cause the vertical deflection on a cathode-ray oscillograph connected at O.

A suitable frequency for comparison could be taken in a similar manner from the ticking of a chronometer, but a still greater degree of accuracy is to be obtained with the aid of a quartz oscillator. To this end use is made of one of the characteristic vibrations of a piezo-electric quartz plate which in the electrical sense behaves as an *LC* circuit of very high quality. In combination with a triode with feedback continuous oscillations with a highly stabilized frequency can be obtained. The extent to which this frequency is stabilized depends, for a given manner of oscillations, upon the dimensions of the quartz plate, the angle of its planes with respect to the axis of the crystal, and upon the temperature. The relative frequency variation due to fluctuations in temperature can be kept at about 10^{-6} per °C 2) by a suitable choice of the

---

1) The damping of an escapement alone is so extremely small that after it is put in motion by a slight touch it continues to run long enough for the test to be made, which, as will appear farther on, takes only a few minutes.

2) The frequency of a quartz oscillator can be accurately determined by comparing it with one of the standardized frequencies transmitted by various radio stations, for instance 200 000 c/s used by Droitwich of the B.B.C., which shows a relative deviation less than 10^{-6}.

cut and the manner of oscillation. This variation corresponds to a gain or loss of no more than one second in 12 days.

The characteristic frequency of a quartz plate of reasonable dimensions (thickness of the order of a few millimeters, diameter of the order of a few centimeters) lies at about 10^4—10^6 c/s, which is much too high for direct comparison with a frequency of 4—6 c/s. By applying the principle of frequency division, however, it is possible to derive from the characteristic frequency a frequency that is a whole number of times lower. Such a division can be brought about in various ways. In the case in question here use has been made of synchronised oscillators with sinusoidal oscillations [3]), the action of which will be briefly considered here.

In order to derive from the frequency of the quartz oscillator, which in our case amounts to 72 900 c/s, for instance a frequency three times as low (24 300 c/s) a simple valve oscillator is adjusted as closely as possible to a characteristic frequency of 24 300 c/s and then coupled in a suitable way to the quartz oscillator. Provided a certain condition that will presently be discussed is satisfied, the synchronism will be maintained — *i.e.* the frequency of the second oscillator will remain exactly 1/3 of the frequency of the quartz oscillator — even if the characteristic frequency of the second oscillator should vary within certain limits (due to temperature changes or fluctuations in the supply voltage).

The effect of the injection of a high-frequency voltage into the oscillator with a lower frequency averages out to nothing if that voltage should be active during the whole cycle of the low frequency. Consequently this has to be avoided, and it is possible to do so in a simple way by applying the high-frequency to the grid of the valve in the low-frequency oscillator and at the same time applying to that grid such a high negative bias as to cause anode current to flow only during a small part of each cycle (*fig. 2*). As a result it is only during these brief intervals that the synchronising oscillation with the higher frequency can make its effect felt (provided suitable amplitudes are chosen). The synchronising voltage is then active only once in each p cycles (by division p : 1) of that frequency, thus only then contributing to the supply of energy to the oscillator circuit which is tuned to the lower frequency.

If it is desired to give a certain ratio of division p strong preference over neighbouring ratios (p—1 and $p + 1$) then a limit has to be set for the order of p. If the frequency f_0 of the quartz oscillator were divided by a factor of for instance $p = 10$ then the risk of the characteristic frequency of the second oscillator varying from the nominal value $f_0/10$ to a neighbouring frequency fraction $f_0/9$ or $f_0/11$ and the oscillation frequency jumping to one of these values is much greater than when one chooses for instance $p = 3$, where the adjacent frequency fractions are relatively farther removed from the nominal value. A high value of the ratio p therefore makes high demands as regards stability of the component parts of the circuit and of the supply voltage.

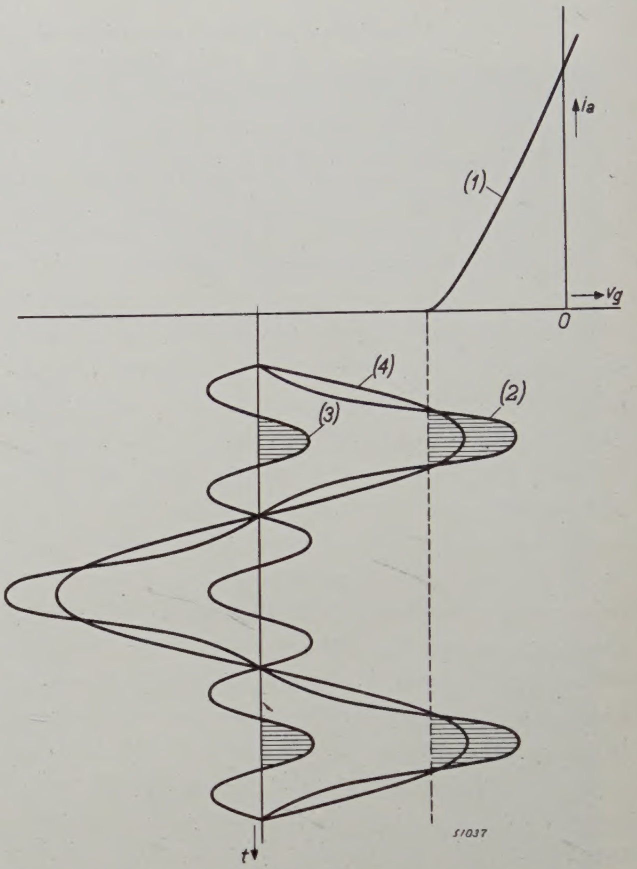


Fig. 2. Curve (*1*) is the anode current i_a as a function of the grid voltage v_g of the oscillator valve in an oscillating circuit the frequency of which must be p times smaller than a control frequency f_0 (here $p = 3$). The grid alternating voltage (*2*), plotted as a function of the time t, consists of two components (*3*) and (*4*) with frequencies of f_0 and f_0/p respectively. The adjustment of the valve and the amplitudes of (*3*) and (*4*) are so chosen that an anode current flows only during a small part of the cycle of (*4*). Only one of each p positive half-waves of (*3*) contributes to the anode current, so that the oscillator is excited with a frequency f_0/p.

Since in our case any jump of the frequency would lead to entirely erroneous conclusions any possibility of this occurring must be precluded. For this reason p is chosen not higher than 5. By connecting a number of dividing stages in cascade an arbitrarily low frequency can also be obtained with small values of p.

---

[3]) I. Koga, A new frequency transformer or frequency changer, Proc. I.R.E. **15**, 669-678, 1927.

What then is the most suitable value for the standard frequency? As already stated, for the escapements of watches three frequencies have been standardized: 4,5 and 6 c/s. Obviously it is preferable that only one frequency should be needed as comparison for these standardized frequencies. Therefore it should be a multiple of 4,5 and 6, thus say 60 c/s. With the experimental apparatus described here the standard frequency for comparison is in fact 60 c/s. It is derived by frequency division in six stages from the frequency of 72 900 c/s supplied by the quartz oscillator. The principle of the set-up is indicated in broad lines in *fig. 3*, while a photograph of the arrangement is given in *fig. 4*.

The time-base voltage of a normal cathode-ray oscillograph is synchronized with this standard frequency, whilst the impulse of the ticking of the timepiece under test provides the vertical deflection, as is to be seen in *fig. 5*. When, therefore, a watch with the nominal frequency of 5 ticks per second is tested each tick is recorded on every twelfth time-base line described on the screen. If the rate of the timepiece is correct then the image of the tick impulse will appear exactly at the same place every time, whereas if the watch regularly gains or loses then the image shifts regularly to the left or to the right, whilst if the rate is irregular this will be reflected in an irregular displacement along the time-base line.

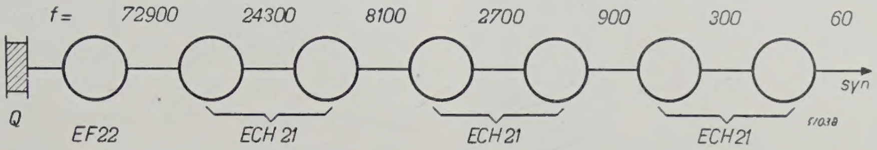


Fig. 3. By means of frequency division one may derive from the frequency 72900 c/s of a quartz plate Q, oscillating by means of a pentode EF 22 with feedback, the frequency of 60 c/s with which the time-base voltage of a normal cathode-ray oscillograph is synchronized. In this instance the frequency division takes place in five stages with triple division and one stage with quintuplicate division. The numbers in the diagram indicate the frequencies f in the various stages in cycles per second. By using valves of the ECH 21 type, containing two electrode systems (a triode and a heptode), the total number of amplifying valves required to effect the frequency division 72900 : 60 can be reduced to three.

Let us now consider the case where a timepiece (with a nominal frequency of n ticks per second) has a small error of say 5 seconds per 24 hours. How much time will it then take to produce with this apparatus an easily discernible displacement of the impulse on the oscillograph screen, for instance over the whole length of the time-base?

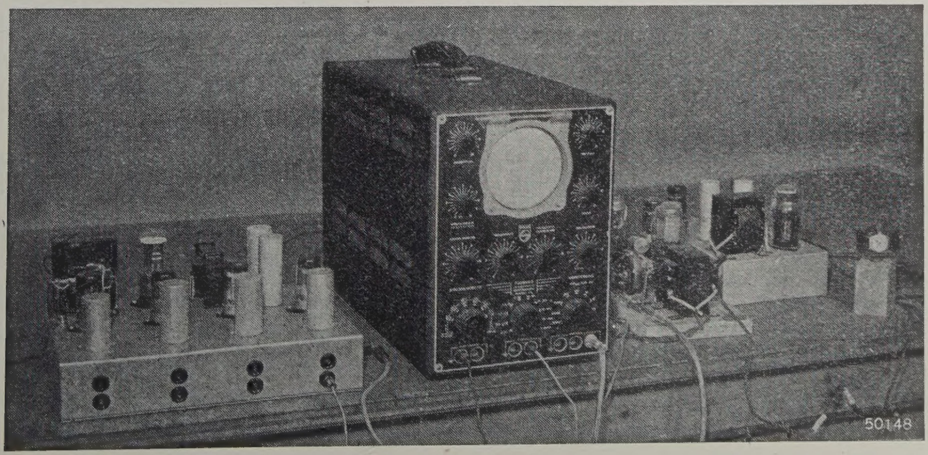


Fig. 4. Experimental arrangement of the apparatus for measuring the rate of watches. In the centre the cathode-ray oscillograph, to the left the apparatus supplying the synchronizing voltage with a frequency of 60 c/s. (see fig. 3). On the extreme right the microphone with a watch placed on top. Between the microphone and the oscillograph are shown the microphone amplifier and the apparatus for converting the ticks of the watch into voltage impulses (see fig. 1).

In 24 hours there are 86 400 seconds. The presumed error in the rate of the timepiece is therefore 5 : 86 400. Between two successive ticks the length of the time-base should be described exactly $60/n$ times. The displacement of one impulse with respect to the previous one is therefore $(5/86\,400) \times (60/n)$ times the length of the time-base. This displacement occurs in $1/n$ second (the interval between two ticks). The displacement of the impulse grows to the length of the time-base in

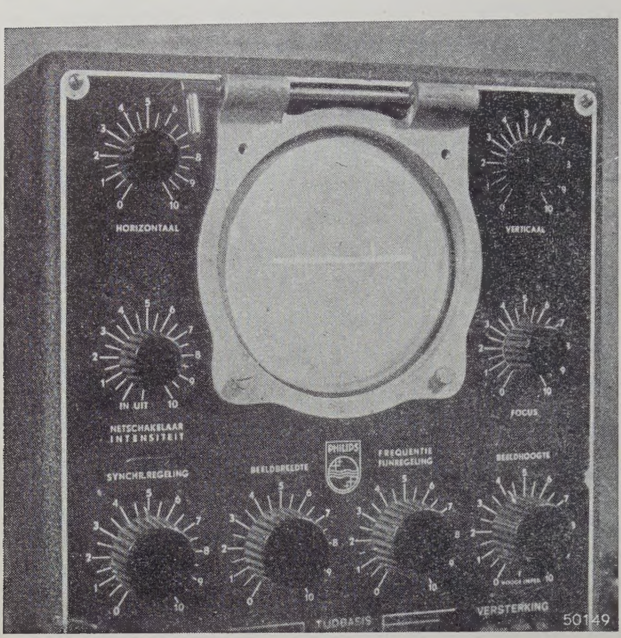


Fig. 5. The voltage impulses generated by the ticking of the watch are visible on the screen of the oscillograph in the shape of peaks, which if the watch is keeping correct time appear at the same place every time but shift when the watch is gaining or losing. The speed at which the peaks shift is a measure for the deviation from the correct time-keeping. Even a small deviation can already be measured in a few minutes.

$$\frac{86\,400}{5} \cdot \frac{n}{60} \cdot \frac{1}{n} \text{ sec} = 288 \text{ sec} \approx 5 \text{ min.}$$

Since it is not necessary to wait until the impulse has been displaced over the entire length of the time-base, the error mentioned can easily be observed within a few minutes. The fact that only such a short time is needed is due to the standard frequency being 10—15 times as high as the ticking frequency, for this means that the time scale is extended as it were by a factor of 10—15, thus showing the displacement all the sooner. Even a complete series of tests with the timepiece in different positions or fully wound up or almost run down takes only a relatively short time. It may be remarked here that a watch with the presumed error of only 5 seconds in 24 hours is regarded as a very good time-keeper. A timepiece running less accurately shows a certain displacement of the impulse in a correspondingly shorter time.